

201359

STIC-Biotech/ChemLib

From: Ramirez, Delia
Sent: Tuesday, September 12, 2006 12:19 PM
To: STIC-Biotech/ChemLib
Subject: 10/806370

CRFE

Hi,

I would like to request the following searches:

1. SEQ ID NO: 11 and 12 in the protein databases (commercial and interference).
2. SEQ ID NO:11 in the nucleic acid databases (commercial and interference).

Please provide a printout of the results.

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

78701

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 07:43:56 ; Search time 9692 Seconds
(without alignments)
11533.233 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 1748
Sequence: 1 ctcagagaccatgagagag99.....gactctgctgacccagagaa 1748

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_env: *
2: gb_pat: *
3: gb_ph: *
4: gb_pl: *
5: gb_pr: *
6: gb_ro: *
7: gb_scs: *
8: gb_sy: *
9: gb_un: *
10: gb_vi: *
11: gb_ov: *
12: gb_hcg: *
13: gb_in: *
14: gb_om: *
15: gb_da: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1 | 1748 | 100.0 | 1748 | 2 | ARS34544 |
| 2 | 1748 | 100.0 | 1748 | 2 | AX369323 |
| 3 | 1687.4 | 96.5 | 1689 | 2 | CS075405 |
| 4 | 1687.4 | 96.5 | 1689 | 2 | DD194635 |
| 5 | 1687.4 | 96.5 | 1689 | 2 | AX149579 |
| 6 | 1670.4 | 95.6 | 2380 | 2 | CS081489 |
| 7 | 1670.4 | 95.6 | 2393 | 2 | DD200135 |
| 8 | 1670.4 | 95.6 | 2393 | 2 | AB048796 |
| 9 | 1670.2 | 95.5 | 2237 | 5 | AY190317 |
| 10 | 1663.4 | 95.2 | 2067 | 2 | CS075313 |
| 11 | 1663.4 | 95.2 | 2067 | 2 | DD194636 |
| 12 | 1663.4 | 95.2 | 2067 | 2 | AX149581 |
| 13 | 1523.4 | 87.2 | 3222 | 5 | BC018715 |
| 14 | 1479 | 84.6 | 3324 | 5 | AB048797 |
| 15 | 1429 | 81.8 | 1614 | 2 | AX360089 |
| 16 | 1302.8 | 74.5 | 1314 | 2 | AX190543 |
| 17 | 1302.8 | 74.5 | 1314 | 2 | AX190546 |
| 18 | 1277.8 | 73.1 | 1574 | 2 | AX704683 |

| | | | | | | |
|----|--------|------|--------|---|----------|--------------------|
| 19 | 1277 | 73.1 | 1375 | 2 | CQ729940 | CQ729940 Sequence |
| 20 | 1258.6 | 72.0 | 1434 | 2 | ARS83622 | ARS83622 Sequence |
| 21 | 1252.8 | 71.7 | 2192 | 2 | BD156842 | BD156842 Primer fo |
| 22 | 1252.8 | 71.7 | 2192 | 2 | AX877776 | AX877776 Sequence |
| 23 | 1252.8 | 71.7 | 2192 | 5 | AK027798 | AK027798 Homo sapi |
| 24 | 1168 | 66.8 | 3320 | 6 | BC085323 | BC085323 Mus muscu |
| 25 | 1111.2 | 63.6 | 2992 | 6 | BC042878 | BC042878 Mus muscu |
| 26 | 1068.4 | 61.1 | 1078 | 2 | AX190545 | AX190545 Sequence |
| 27 | 1062.6 | 60.2 | 1341 | 2 | ARS83623 | ARS83623 Sequence |
| 28 | 1053 | 60.2 | 1230 | 2 | ARS34553 | ARS34553 Sequence |
| 29 | 1053 | 60.2 | 1230 | 2 | AX369347 | AX369347 Sequence |
| 30 | 782.4 | 44.8 | 1222 | 2 | BD276278 | BD276278 Serine pr |
| 31 | 709.4 | 40.6 | 834 | 2 | ARS34546 | ARS34546 Sequence |
| 32 | 709.4 | 40.6 | 834 | 2 | AX369340 | AX369340 Sequence |
| 33 | 535.6 | 30.6 | 852 | 2 | BD149569 | BD149569 Primer fo |
| 34 | 535.6 | 30.6 | 852 | 2 | AX869507 | AX869507 Sequence |
| 35 | 446 | 25.5 | 2125 | 6 | BC010843 | BC010843 Mus muscu |
| 36 | 440.8 | 25.2 | 678 | 2 | ARS34548 | ARS34548 Sequence |
| 37 | 440.8 | 25.2 | 678 | 2 | AX369342 | AX369342 Sequence |
| 38 | 411.6 | 23.5 | 614 | 2 | ARS34551 | ARS34551 Sequence |
| 39 | 411.6 | 23.5 | 614 | 2 | AX369345 | AX369345 Sequence |
| 40 | 394.8 | 22.6 | 160222 | 5 | AP002962 | AP002962 Homo sapi |
| 41 | 394.8 | 22.6 | 184926 | 5 | AP000757 | AP000757 Homo sapi |
| 42 | 348 | 19.9 | 402 | 2 | ARS34534 | ARS34534 Sequence |
| 43 | 348 | 19.9 | 402 | 2 | AX369313 | AX369313 Sequence |
| 44 | 335.8 | 19.2 | 621 | 2 | ARS34547 | ARS34547 Sequence |
| 45 | 335.8 | 19.2 | 621 | 2 | AX369341 | AX369341 Sequence |

ALIGNMENTS

| | | | | | |
|----------------------------|---|--|-----|--------|-----------------|
| RESULT 1 | ARS34544 | 1748 bp | DNA | linear | PAT 08-OCT-2004 |
| LOCUS | ARS34544 | | | | |
| DEFINITION | Sequence 11 from patent US 6734006. | | | | |
| ACCESSION | ARS34544 | | | | |
| VERSION | ARS34544.1 | GI:53924854 | | | |
| KEYWORDS | | | | | |
| SOURCE | Unknown. | | | | |
| ORGANISM | Unknown. | | | | |
| REFERENCE | 1 (bases 1 to 1748) | | | | |
| AUTHORS | Xiao, Y. and Gedrich, R.W. | | | | |
| TITLE | Regulation of human transmembrane serine protease | | | | |
| JOURNAL | Patent: US 6734006-A, 11 11-MAY-2004; | | | | |
| | Bayer Aktiengesellschaft; Leverkusen; | | | | |
| | WOX; | | | | |
| FEATURES | | | | | |
| source | Location/Qualifiers | | | | |
| | 1..1748 | | | | |
| | /organism="unknown" | | | | |
| | /mol_type="genomic DNA" | | | | |
| ORIGIN | | | | | |
| Query Match | 100.0%; Score 1748; DB 2; Length 1748; | | | | |
| Best Local Similarity | 100.0%; Pred. No. 0; | | | | |
| Matches 1748; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| QY | 1 | CTCAGAGACCATGAGAGGACGACCGGAGATGATCTTCACAGCAAGAACCTTCAGC | 60 | | |
| DB | 1 | CTCAGAGACCATGAGAGGACGACCGGAGATGATCTTCACAGCAAGAACCTTCAGC | 60 | | |
| QY | 61 | TGAGGATCTTCAGCCAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC | 120 | | |
| DB | 61 | TGAGGATCTTCAGCCAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC | 120 | | |
| QY | 121 | CCAGGATCTTCAGCCAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC | 180 | | |
| DB | 121 | CCAGGATCTTCAGCCAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC | 180 | | |
| QY | 181 | CCAGGATCTTCAGCTGGTACCTTCAGGCGGGGATCTTCAGC | 240 | | |
| DB | 181 | CCAGGATCTTCAGCTGGTACCTTCAGGCGGGGATCTTCAGC | 240 | | |

| | | | |
|----|------|--|------|
| OY | 241 | CCAGGCATCTCCAGGCCCGGGGACATCCCGGGCTCTGGGACATCACTTTCCAGTCTCTACATCCGG | 300 |
| Db | 241 | CCAGGCATCTCCAGGCCCGGGGACATCTCCGGCTCTGGGACATCACTTTCCAGTCTCTACATCCGG | 300 |
| OY | 301 | CAGGTCATCATCCGCGCAGGTCAGACCTCGGTTGCAACACTTCCCAACCAAGTGTACTTGT | 360 |
| Db | 301 | CAGGTCATCATCCGCGCAGGTCAGACCTCGGTTGCAACACTTCCCAACCAAGTGTACTTGT | 360 |
| OY | 361 | TAGAGCAACACCAAGTGGGGGCTGTACCCATCCGATCATCTCTCTGCGAGGTGAGCACAGC | 420 |
| Db | 361 | TAGAGCAACACCAAGTGGGGGCTGTACCCATCCGATCATCTCTCTGCGAGGTGAGCACAGC | 420 |
| OY | 421 | AACCGGGGCGCCACAGGGAGAGCGCCAGGTATCCAGAGCTCTGCCAAGTTCACTTGGCGGGAGGG | 480 |
| Db | 421 | AACCGGGGCGCCACAGGGAGAGCGCCAGGTATCCAGAGCTCTGCCAAGTTCACTTGGCGGGAGGG | 480 |
| OY | 481 | CCAGAGAGAGCTACCGGTCATCGGATGGGTGTCTCTCTCATTTGGCCCTGATGTTTCGCT | 540 |
| Db | 481 | CCAGAGAGAGCTACCGGTCATCGGATGGGTGTCTCTCTCATTTGGCCCTGATGTTTCGCT | 540 |
| OY | 541 | CATCATCTCTTTCCAGTTCTTGGCAGAGGGCCACACAGGAAATCAGGTATCAAGAGCAGAGGGA | 600 |
| Db | 541 | CATCATCTCTTTCCAGTTCTTGGCAGAGGGCCACACAGGAAATCAGGTATCAAGAGAGCAGAGGGA | 600 |
| OY | 601 | GAGCGTCCCAAGGACGCGGTTTGGGTGTGACGGGGGTGGATGGACTGGCAAGCTGGAAGAGTGA | 660 |
| Db | 601 | GAGCGTCCCAAGGACGCGGTTTGGGTGTGACGGGGGTGGATGGACTGGCAAGCTGGAAGAGTGA | 660 |
| OY | 661 | CGAGCTGGGCTGCGTAGAGATTGACTGGGACCAAGTCTGTGCTTAATAATCTACTCTGGGCTC | 720 |
| Db | 661 | CGAGCTGGGCTGCGTAGAGATTGACTGGGACCAAGTCTGTGCTTAATAATCTACTCTGGGCTC | 720 |
| OY | 721 | CTCCCATCAGTGGCTTTCCCATTTGTATGACGACACTGGAATGACTCTTACTCAGAGAAAC | 780 |
| Db | 721 | CTCCCATCAGTGGCTTTCCCATTTGTATGACGACACTGGAATGACTCTTACTCAGAGAAAC | 780 |
| OY | 781 | CTGGCAGAGAGCTGGGTTTGGAGAGTGCACCCGGACAACCGAGGTTGGCCACAGGGAATTT | 840 |
| Db | 781 | CTGGCAGAGAGCTGGGTTTGGAGAGTGCACCCGGACAACCGAGGTTGGCCACAGGGAATTT | 840 |
| OY | 841 | TGCCAACAGCTTCTGCATCTTGAGATTAACAATCCACCATCCAGAGAAAGCTTCCACAGTTC | 900 |
| Db | 841 | TGCCAACAGCTTCTGCATCTTGAGATTAACAATCCACCATCCAGAGAAAGCTTCCACAGTTC | 900 |
| OY | 901 | TGAATGCCCTTCCAGCGGTAATATCTCCCTCCAGTGTTCACATGCGGACCTGAGGGCCAT | 960 |
| Db | 901 | TGAATGCCCTTCCAGCGGTAATATCTCCCTCCAGTGTTCACATGCGGACCTGAGGGCCAT | 960 |
| OY | 961 | GACCGGGGAGATCTGGTGGAGAGGGGGCGCTGGCTCCGATATGACAATGGGCTTGGCAAGTAG | 1020 |
| Db | 961 | GACCGGGGAGATCTGGTGGAGAGGGGGCGCTGGCTCCGATATGACAATGGGCTTGGCAAGTAG | 1020 |
| OY | 1021 | TCTGACTTTCGGCACACCCACATCTGTGAGGACGCTCATATTGACGCGCCACAGTGGGTTCT | 1080 |
| Db | 1021 | TCTGACTTTCGGCACACCCACATCTGTGAGGACGCTCATATTGACGCGCCACAGTGGGTTCT | 1080 |
| OY | 1081 | CACCTGCGGCCACATGCTCTTCTGTGACCCCGGAGAAAGGTCTTGGAGGCTGGAAGTGTGA | 1140 |
| Db | 1081 | CACCTGCGGCCACATGCTCTTCTGTGACCCCGGAGAAAGGTCTTGGAGGCTGGAAGTGTGA | 1140 |
| OY | 1141 | CGCGGGGACACAGCAACTGAGACCAATGTCGTGAGGACAGCTTCCATTTGGCGAGATCATAT | 1200 |
| Db | 1141 | CGCGGGGACACAGCAACTGAGACCAATGTCGTGAGGACAGCTTCCATTTGGCGAGATCATAT | 1200 |
| OY | 1201 | CAACGACATTAACACCGAATGAGAGAGAGCACTTATGACATGCGCTCATAGCGGCTGTCCAA | 1260 |
| Db | 1201 | CAACGACATTAACACCGAATGAGAGAGAGCACTTATGACATGCGCTCATAGCGGCTGTCCAA | 1260 |
| OY | 1261 | GCCCCCTGACCTGTCCGCTCAGATCAATCAACCTGTGCTCTCCCAAGCATGAGCAGACCTT | 1320 |
| Db | 1261 | GCCCCCTGACCTGTCCGCTCAGATCAATCAACCTGTGCTCTCCCAAGCATGAGCAGACCTT | 1320 |

| | | | |
|-----------------------|---|--|-----------------|
| Qy | 1321 | TAGCCTCAATGAGACCTGTGCGATCACAGGCTTTGGCAAGCCAGGGAGACAGATGACAA | 1380 |
| Db | 1321 | TAGCCTCAATGAGACCTGTGCGATCACAGGCTTTGGCAAGCCAGGGAGACAGATGACAA | 1380 |
| Qy | 1381 | GACATCCCTCTCTCCGGAGGTGCAGGTCAATCTGATCGACTTCAAGAAATGCAATGA | 1440 |
| Db | 1381 | GACATCCCTCTCTCCGGAGGTGCAGGTCAATCTGATCGACTTCAAGAAATGCAATGA | 1440 |
| Qy | 1441 | CTACTTGGTCTATATACAGTTAACCCTTACCAGGATATATGTGTGTGGGACCTTGCTGG | 1500 |
| Db | 1441 | CTACTTGGTCTATATACAGTTAACCCTTACCAGGATATATGTGTGTGGGACCTTGCTGG | 1500 |
| Qy | 1501 | GGGACAGACACTCTGCAGGAGAGACAGCGGGGGGCTCTTGCTGTGAGACAGAACACCG | 1560 |
| Db | 1501 | GGGACAGACACTCTGCAGGAGAGACAGCGGGGGGCTCTTGCTGTGAGACAGAACACCG | 1560 |
| Qy | 1561 | CTGTATCTTGCGACAGGTATCACAGCTGGGGACAGGCTGTGGCCAGAGAAACAAACCTGG | 1620 |
| Db | 1561 | CTGTATCTTGCGACAGGTATCACAGCTGGGGACAGGCTGTGGCCAGAGAAACAAACCTGG | 1620 |
| Qy | 1621 | TGTGTATACCCAAAGATGACAGAAAGTTCTTCCCTGGATTATACAGCAAGATGAGAGCAGGT | 1680 |
| Db | 1621 | TGTGTATACCCAAAGATGACAGAAAGTTCTTCCCTGGATTATACAGCAAGATGAGAGCAGGT | 1680 |
| Qy | 1681 | GCGATTGAGAAAATCTTAACCAAGCTGGCTGTGCTGTGACACAGACCGGCTGCTGTGAC | 1740 |
| Db | 1681 | GCGATTGAGAAAATCTTAACCAAGCTGGCTGTGCTGTGACACAGACCGGCTGCTGTGAC | 1740 |
| Qy | 1741 | TCGAGAAA | 1748 |
| Db | 1741 | TCGAGAAA | 1748 |
| RESULT 2 | | | |
| AX369323 | AX369323 | 1748 bp | DNA |
| LOCUS | AX369323 | 1748 bp | DNA |
| DEFINITION | Sequence 11 from Patent WO0196538. | linear | PAT 16-FEB-2002 |
| ACCESSION | AX369323 | | |
| VERSION | AX369323.1 | GI:18857263 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | Homo sapiens (human) | | |
| REFERENCE | 1. Xiao, Y. and Gedrich, R. | | |
| AUTHORS | Regulation of human transmembrane serine protease | | |
| TITLE | Patent: WO 0196538-A 11 20-DEC-2001; | | |
| JOURNAL | Bayer Aktiengesellschaft (DE) | | |
| FEATURES | Location/Qualifiers | | |
| source | 1. 1748 | | |
| ORIGIN | /organism="Homo sapiens" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9606" | | |
| Query Match | 100.0%; | Score 1748; | DB 2; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | |
| Matches 1748; | Conservative 0; | Mismatches 0; | Indels 0; |
| Gaps 0; | | | |
| Qy | 1 | CTCAGAGACCATGAGAGGGACACGCCACCGGAATGCATCTTCCAGCAAGAACACACTTCAGC | 60 |
| Db | 1 | CTCAGAGACCATGAGAGGGACACGCCACCGGAATGCATCTTCCAGCAAGAACACACTTCAGC | 60 |
| Qy | 61 | TGAGGCAATCTTCAGCCAGGACATCTTCAGCTGGGACACTTCAGGCCGGGCAATCTCCAGC | 120 |
| Db | 61 | TGAGGCAATCTTCAGCCAGGACATCTTCAGCTGGGACACTTCAGGCCGGGCAATCTCCAGC | 120 |
| Qy | 121 | CCAGGCAATCTTCAGCCAGGACATCTTCAGCTGGGACACTTCGGGCCGGGCAATCTCCAGC | 180 |
| Db | 121 | CCAGGCAATCTTCAGCCAGGACATCTTCAGCTGGGACACTTCGGGCCGGGCAATCTCCAGC | 180 |


```

Db      121  CCAGCCCGGATCTCCAGCTGGAGCAACCTCCGGCCGGGATCTCCAGCCCAAGCATCT
Qy      191  CCAGCTGGTACACTCCAGGCGGGGCACTCCAGCCGGGATCTCCAGCCCAAGCATCT
Db      181  CCAGCTGGTACACTCCAGGCGGGGCACTCCAGGCGGGGATCTCCAGCCCAAGCATCT
Qy      251  CCAGCCCGGATCTCCAGGCTCTGGCATCACTTTCAGAGTCTGATCCGGAGGATCACTCA
Db      241  CCAGCCCGGATCTCCAGGCTCTGGCATCACTTTCAGAGTCTGATCCGGAGGATCACTCA
Qy      311  TCCGCGAGGTGAGCTCGGTGACAACTCCCAACAGAGTGTACTTTAGACAAACA
Db      301  TCCGCGAGGTGAGCTCGGTGACAACTCCCAACAGAGTGTACTTTAGAGCAACA
Qy      371  CCAGTGGGGGGGTATACCATCCGATCATCTCTGCGAGGTGAGAGCAACAGAGGCC
Db      361  CCAGTGGGGGGGTATACCATCCGATCATCTCTGCGAGGTGAGAGCAACAGAGGCC
Qy      431  ACCAGGGAGAGCCCAAGGTACAGAGCTGCGCAAGTTCACTTGGCGGGAGGCGAGAGCAG
Db      421  ACCAGGGAGAGCCCAAGGTACAGAGCTGCGCAAGTTCACTTGGCGGGAGGCGAGAGCAG
Qy      491  CTACCGCTCATCGGGTGGGTGCTCTCTCTCATTTGCTGCTGTTGCTCATCATCTCTC
Db      481  CTACCGCTCATCGGGTGGGTGCTCTCTCTCATTTGCTGCTGTTGCTCATCATCTCTC
Qy      551  TTCCAGTTCTGGAGGGGCAACAGAGGATCAGGTACAGAGGAGAGAGAGGCTGTGCC
Db      541  TTCCAGTTCTGGAGGGGCAACAGAGGATCAGGTACAGAGGAGAGAGAGGCTGTGCC
Qy      611  AAGCAGCTGTCTGCTGTGACGGGGGTGGTGGAGCTGCAAGCTGAAAGTGAAGAGCTGGC
Db      601  AAGCAGCTGTCTGCTGTGACGGGGGTGGTGGAGCTGCAAGCTGAAAGTGAAGAGCTGGC
Qy      671  TGCGTGAAGTTTGACTGGGACAAAGTCTTGTAAATCTACTGGGTCTTCCCATCAG
Db      661  TGCGTGAAGTTTGACTGGGACAAAGTCTTGTAAATCTACTGGGTCTTCCCATCAG
Qy      731  TGCGTCCCATCTGTAGAGCAAGCTGGAATGACTCTACTAGAGAGAGAGCTGCGAGAG
Db      721  TGCGTCCCATCTGTAGAGCAAGCTGGAATGACTCTACTAGAGAGAGAGCTGCGAGAG
Qy      791  CTGGGTTTCAGAGAGTGTCTCACCGGACAAACAGAGTTGGCCACAGGATTTTTCACAGC
Db      781  CTGGGTTTCAGAGAGTGTCTCACCGGACAAACAGAGTTGGCCACAGGATTTTTCACAGC
Qy      851  TTCTCAATCTTGAATACAACTCCACCATCCAGAGAAAGCTTCCACAGGTCTGAATGCCCT
Db      841  TTCTCAATCTTGAATACAACTCCACCATCCAGAGAAAGCTTCCACAGGTCTGAATGCCCT
Qy      911  TCCGAGCGGTATATCTCCCTCCAGGTGTCCCACTGCGGATGAGAGGCGCATGACCGGGCGG
Db      901  TCCGAGCGGTATATCTCCCTCCAGGTGTCCCACTGCGGATGAGAGGCGCATGACCGGGCGG
Qy      971  ATCGTGGAGGGGCGCTGGCTCGGATAGCAAGTGCGCTTGGCAAGTGAAGTCTGCACCTTC
Db      961  ATCGTGGAGGGGCGCTGGCTCGGATAGCAAGTGCGCTTGGCAAGTGAAGTCTGCACCTTC
Qy      1031  GGCACCAACCACTCTGTGAGAGGACCGCTCATTTGACGCCAGTGGGTCTCACTGCGGCC
Db      1021  GGCACCAACCACTCTGTGAGAGGACCGCTCATTTGACGCCAGTGGGTCTCACTGCGGCC
Qy      1091  CACTGCTTCTTCTGTGACCGGGGAGAGGTTCTGGAGGGCTGGAAGGTGTACGCGGGCACCC
Db      1081  CACTGCTTCTTCTGTGACCGGGGAGAGGTTCTGGAGGGCTGGAAGGTGTACGCGGGCACCC
Qy      1151  AGCAACCTGACCAAGTTGCTGAGGAGAGCTTCAATTGCGGAGATCATCATCAAGCAAT
Db      1141  AGCAACCTGACCAAGTTGCTGAGGAGAGCTTCAATTGCGGAGATCATCATCAAGCAAT
Qy      1211  TACACCGATGAGAGAGAGATATGACATGGCTCATGGGCTGTCTCAAGCCCTTGACC

```

```

Db      1201  TACACCGATGAGAGAGAGATATGACATGGCCCTCATGGGCTGTCTCAAGCCCTTGACC
Qy      1271  CTGTGCGCTGACATCCACCGTGGCTGCTCCCATGATGAGAGACCTTTAGCTCAAT
Db      1261  CTGTGCGCTGACATCCACCGTGGCTGCTCCCATGATGAGAGACCTTTAGCTCAAT
Qy      1331  GAGACCTGTGATACAGAGCTTTGGCAAGACAGAGGAGACAGATGACAAAGATCTCCCC
Db      1321  GAGACCTGTGATACAGAGCTTTGGCAAGACAGAGGAGACAGATGACAAAGATCTCCCC
Qy      1391  TTCTCCGGGAGGTGACAGGTTCATCTGATGACTTCAAGAAATGCAATGACTTCTGTC
Db      1381  TTCTCCGGGAGGTGACAGGTTCATCTGATGACTTCAAGAAATGCAATGACTTCTGTC
Qy      1451  TATGACAGTTACCTTACCCCAAGATGATGTGTCTGGAGACCTTGTGGGGGAGAGAC
Db      1441  TATGACAGTTACCTTACCCCAAGATGATGTGTCTGGAGACCTTGTGGGGGAGAGAC
Qy      1511  TCTGCCAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGAGAAACAACGCTGGTACCTG
Db      1501  TCTGCCAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGAGAAACAACGCTGGTACCTG
Qy      1571  GCAGGTGTACACAGCTGGGGGACAGAGCTGTGGCCAGAGAAACAACCTGGTGTACACC
Db      1561  GCAGGTGTACACAGCTGGGGGACAGAGCTGTGGCCAGAGAAACAACCTGGTGTACACC
Qy      1631  AAAGTGACAGAGTTCTTCCCTGGATTTTACAGCAAGATGAGAGAGGATGCAATTGAGA
Db      1621  AAAGTGACAGAGTTCTTCCCTGGATTTTACAGCAAGATGAGAGAGGATGCAATTGAGA
Qy      1691  AAATCTTAA 1699
Db      1681  AAATCTTAA 1689

RESULT 4
DD194635
LOCUS
DEFINITION
DD194635
1689 bp DNA linear PAT 19-JAN-2006
Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
Therof.
DD194635
VERSION
DD194635.1 GI:85634726
KEYWORDS
JP 2005518332-A/12.
SOURCE
unidentified
ORGANISM
unidentified
REFERENCE
1 (bases 1 to 1689)
AUTHORS
Komandla,M., Sempke,J.E., Vlaeuk,G.P., Kemp,S.J., Siev,D.V. and
Madison,E.L.
TITLE
Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
Therof.
JOURNAL
Patent: JP 2005518332-A 12 23-JUN-2005;
Edwin I. Madison, Joseph Edward Sempke, George P Vlaeuk, Scott Jeffrey
Kemp, Mallareddy Komandla, Daniel Vanna Siev
COMMENT
PN JP 2005518332-A/12
PD 23-JUN-2005
PR 23-MAY-2002 JP 2002592470
PR 23-MAY-2001 US 60/293267
PI mallareddy komandla, joseph edward sempke, george p vlaeuk, pi
scott jeffrey kemp, daniel vanna siev, edwin i madison CC
FH Key Location/Qualifiers
FT CDS (1) . . . (1689)
FT /note="Nucleic acid encoding Endothelase 2-s
protein".
FEATURES
source Location/Qualifiers
1. .1689
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 96.5%; Score 1687.4; DB 2; Length 1689;

```

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 11 ATGAGAGGGGACACCCAGGAGATGATCTCTCCAGCAGAGAACACTTTCAGCTGGAGCATCT 70
Db 1 ATGAGAGGGGACACCCAGGAGATGATCTCTCCAGCAGAGAACACTTTCAGCTGGAGCATCT 60
QY 71 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 130
Db 61 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 120
QY 131 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 190
Db 121 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 180
QY 191 CCAGCTGTACACTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCGAGGATCT 250
Db 181 CCAGCTGTACACTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCGAGGATCT 240
QY 251 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 310
Db 241 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 300
QY 311 TCCGCAAGCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCTGGGACACCA 370
Db 301 TCCGCAAGCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCTGGGACACCA 360
QY 371 CCAGTGGGGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGGCGGGGATCT 430
Db 361 CCAGTGGGGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGGCGGGGATCT 420
QY 431 ACCAGGAGAGCCGAGTACAGGCTGGGACACCTCCAGGCGGGGATCTCCAGGCGGGGATCT 490
Db 421 ACCAGGAGAGCCGAGTACAGGCTGGGACACCTCCAGGCGGGGATCTCCAGGCGGGGATCT 480
QY 491 CTACCGCTCAGCTGGGATGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 550
Db 481 CTACCGCTCAGCTGGGATGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 551 TTCCAGTTCTGGGACGGGACACAGGATCAGGTACAGGACAGAGGAGAGCTGTGCC 610
Db 541 TTCCAGTTCTGGGACGGGACACAGGATCAGGTACAGGACAGAGGAGAGCTGTGCC 600
QY 611 AAGAGCGCTGTGGCTGGACGGGCTGGGATCTGAGCTGAGGAGAGGATGAGGAGCTGGGC 670
Db 601 AAGAGCGCTGTGGCTGGACGGGCTGGGATCTGAGCTGAGGAGAGGATGAGGAGCTGGGC 660
QY 671 TGGCTGAGGTTTGAAGTCAAGTCTCTGTTAAATCTACTGAGGCTCTCCATCAG 730
Db 661 TGGCTGAGGTTTGAAGTCAAGTCTCTGTTAAATCTACTGAGGCTCTCCATCAG 720
QY 731 TGGCTTCCCATCTGTAGCAGGACAGTGAATGACTCTACTGAGGAGAGAGCTGTCCAGCAG 790
Db 721 TGGCTTCCCATCTGTAGCAGGACAGTGAATGACTCTACTGAGGAGAGAGCTGTCCAGCAG 780
QY 791 CTGGGTTTCCAGAGTGTCTACCCGAGCAACGAGGTTGGCCCAAGGATTTTGGCCAAACAGC 850
Db 781 CTGGGTTTCCAGAGTGTCTACCCGAGCAACGAGGTTGGCCCAAGGATTTTGGCCAAACAGC 840
QY 851 TTTCATCTTGAATACAACTCCACATCCAGGAGAAAGCTCCACAGGCTGAGATGCCCT 910
Db 841 TTTCATCTTGAATACAACTCCACATCCAGGAGAAAGCTCCACAGGCTGAGATGCCCT 900
QY 911 TCCAGCGGATATCTCTCTCCAGTGTTCCTCACTGCGGATGAGGCGCATGACCGGCGG 970
Db 901 TCCAGCGGATATCTCTCTCCAGTGTTCCTCACTGCGGATGAGGCGCATGACCGGCGG 960
QY 971 ATGCTGGAGGGGGGCTGGGCTCGGATAGCAAGTGGGCTTGGCAAGTGTGCACTTC 1030
Db 961 ATGCTGGAGGGGGGCTGGGCTCGGATAGCAAGTGGGCTTGGCAAGTGTGCACTTC 1020
QY 1031 GGACCAACCAATCTGTGAGGACGCTCATTTGACGCGGCTGAGGCTGCTCATCTGCCGCC 1090

```

```

Db 1021 GGACCAACCAATCTGTGAGGACGCTCATTTGACGCCAGTGGGTCTCATCTGCCCC 1080
QY 1091 CACTGCTTCTTGGACCCCGGAGAAAGTCTCTGAGGGCTGGAAAGTGTACGGGGGACC 1150
Db 1081 CACTGCTTCTTGGACCCCGGAGAAAGTCTCTGAGGGCTGGAAAGTGTACGGGGGACC 1140
QY 1151 AGCAACCTGACACGATTTGAGGAGGAGCTCTCAATTCAGGATCATCATCAAGCAAT 1210
Db 1141 AGCAACCTGACACGATTTGAGGAGGAGCTCTCAATTCAGGATCATCATCAAGCAAT 1200
QY 1211 TACACGATGAGAGAGCACTATGATGATGCTCTCAATTCAGGCTGTCCAAAGCCCTGACC 1270
Db 1201 TACACGATGAGAGAGCACTATGATGATGCTCTCAATTCAGGCTGTCCAAAGCCCTGACC 1260
QY 1271 CTGTCGCTCAGATCCACCTGCTGGCTCCCAATGATGAGACAGACTTTAGCTCAAT 1330
Db 1261 CTGTCGCTCAGATCCACCTGCTGGCTCCCAATGATGAGACAGACTTTAGCTCAAT 1320
QY 1331 GAGACCTGTGATCAAGGCTTTGGCAAGACAGGAGACAGATGACAAAGATCCCC 1390
Db 1321 GAGACCTGTGATCAAGGCTTTGGCAAGACAGGAGACAGATGACAAAGATCCCC 1380
QY 1391 TTCTCCGGAGAGTGCAGTCAATCTCATGACTTCAAGAAATGCAATGACTTGGTC 1450
Db 1381 TTCTCCGGAGAGTGCAGTCAATCTCATGACTTCAAGAAATGCAATGACTTGGTC 1440
QY 1451 TATGACGTTACCTTACCCCAAGATATGATGTCTGGGACCTTCTGGGGGAGAGAC 1510
Db 1441 TATGACGTTACCTTACCCCAAGATATGATGTCTGGGACCTTCTGGGGGAGAGAC 1500
QY 1511 TCTGCCAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGCAGAACACCGCTGGTACCTG 1570
Db 1501 TCTGCCAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGCAGAACACCGCTGGTACCTG 1560
QY 1571 GCAAGTGTCAACGAGCTGGGACAGGCTGTGGCCAGAAACAACTGTGTGTACACC 1630
Db 1561 GCAAGTGTCAACGAGCTGGGACAGGCTGTGGCCAGAAACAACTGTGTGTACACC 1620
QY 1631 AAGTGAACAAGTTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGCATTCATA 1690
Db 1621 AAGTGAACAAGTTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGCATTCATA 1680
QY 1691 AAATCTTAA 1699
Db 1681 AAATCTTAA 1689

```

RESULT 5
AX149579 1689 bp DNA linear PAT 31-AUG-2001
AX149579
DEFINITION Sequence 3 from Patent WO0136604.
ACCESSION AX149579
VERSION AX149579.1 GI:14348013
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
1. Madison, E.L. and Ong, E.O.
Nucleic acids encoding endothelases, endothelases and uses
thereof
Patent: WO 0136604-A 3 25-MAY-2001;
CORVAS INTERNATIONAL, INC. (US)
Location/Qualifiers
1. .1689
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .1689
CDS
/note="unnamed protein product; DNA encoding a
transmembrane serine protease (Endothelase 2-s) protein"

```

/codon_start=1
/protein_id="CAC41220.1"
/db_xref="GI:14348014"
/translation="MERDISHGNASPARTPSAGAPAOAPAGTPGRASPAOASPAO
SPACTPGRASPAPAGTPGRASPAOASPAOASPAOASPAOASPAOASPAOASPAO
RBAATVTSPTVTVLVRATPVAVPIRSSPAPASAPATATRSPTSLPKPTWREGOKO
LPLGCVLKLALVSLILFQPMOHTGIRKEORSNNDSEKTCQQLFESNARTTEVAHRD
LGCVPFMDKSLKTYGSSSHQWLPICSSNNNDSEKTCQQLFESNARTTEVAHRD
FANFSILIRVNSTIOESLHRECEPORYISIQCSHCGRAMTGRIVGALASDSKWP
OVSLHFTTHICGGTLIDQWVLTAAHCFPTTRKLVGMVTVAGTSLHDLPEASAI
AEIINAVTDEBDYDIALMRSLKPLTSLSHIPACIPMHGTFPSLNTGCTWTFPGK
TRBTDDKSPFLREYQVNLIDPKCNLDYLVDSYLTTPRMKAGLREGRCQGDSSG
PLVCEONNRMTLAGVTSMGTGQGRNRPVITKVELEWPIYSMESVRFIKS"

```

ORIGIN

Query Match 96.5%; Score 1687.4; DB 2; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 11 ATGAGAGAGGACAGCCAGGAAATGCATCTCCAGCAAGAACCTTCAGCTGAGCATCT 70
Db 1 ATGAGAGAGGACAGCCAGGAAATGCATCTCCAGCAAGAACCTTCAGCTGAGCATCT 60
QY 71 CCAGCCCAAGGATCTCCAGCTGGGACACCTCCAGCCCGGAGCATCTCCAGCCCAAGCATCT 130
Db 61 CCAGCCCAAGGATCTCCAGCTGGGACACCTCCAGCCCGGAGCATCTCCAGCCCAAGCATCT 120
QY 131 CCAGCCCAAGGATCTCCAGCTGGGACACCTCCAGCCCGGAGCATCTCCAGCCCAAGCATCT 190
Db 121 CCAGCCCAAGGATCTCCAGCTGGGACACCTCCAGCCCGGAGCATCTCCAGCCCAAGCATCT 180
QY 191 CCAGCTGTATACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCAAGCATCT 250
Db 181 CCAGCTGTATACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCAAGCATCT 240
QY 251 CCAGCCCGGAGCATCTCCAGCTGGGACACCTTCAGCTGAGCATCTCCAGCTGAGCATCT 310
Db 241 CCAGCCCGGAGCATCTCCAGCTGGGACACCTTCAGCTGAGCATCTTCAGCTGAGCATCT 300
QY 311 TCCGCCAGAGTCAAGCTCGGTGACAAACCTCCCAACAGAGTGTACTTGTAGAGCAACA 370
Db 301 TCCGCCAGAGTCAAGCTCGGTGACAAACCTCCCAACAGAGTGTACTTGTAGAGCAACA 360
QY 371 CCAGTGGGGGTGTATCCCATCCATCTCTGCGCAGGTGAGGACCAAGCAAGAGGCC 430
Db 361 CCAGTGGGGGTGTATCCCATCCATCTCTGCGCAGGTGAGGACCAAGCAAGAGGCC 420
QY 431 ACCAGGAGAGGCCCAAGTACAGAGCTGCGCAAGTTTCACTTGGCGGAGGACCAAGAGAG 490
Db 421 ACCAGGAGAGGCCCAAGTACAGAGCTGCGCAAGTTTCACTTGGCGGAGGACCAAGAGAG 480
QY 491 CTACCGCTCATCGGAGTGTGCTCTCTCTCATTTGGCCCTGTGTTTGGCTCATCTCTC 550
Db 481 CTACCGCTCATCGGAGTGTGCTCTCTCTCTCATTTGGCCCTGTGTTTGGCTCATCTCTC 540
QY 551 TTCCAGTTCTTGAGAGGACCAAGAGATCAGGTACAAAGAGAGAGAGAGAGTGTCTCC 610
Db 541 TTCCAGTTCTTGAGAGGACCAAGAGATCAGGTACAAAGAGAGAGAGAGAGTGTCTCC 600
QY 611 AAGAGAGCTGTGCTGTGAGAGGGGTGTGAGCTGCAAGCTGGAAGAGTGAAGAGTGTGGC 670
Db 601 AAGAGAGCTGTGCTGTGAGAGGGGTGTGAGCTGCAAGCTGGAAGAGTGAAGAGTGTGGC 660
QY 671 TGCGTGAAGTTGACTGGAGCAAGTCTGTCTTAAATCTACTCTGGGTCTCTCCATCAG 730
Db 661 TGCGTGAAGTTGACTGGAGCAAGTCTGTCTTAAATCTACTCTGGGTCTCTCCATCAG 720
QY 731 TGCGTCTCCATCTGTAGAGCAACTGGAATGACTCTTACAGAGAGAGAGTGTCCAGAG 790
Db 721 TGCGTCTCCATCTGTAGAGCAACTGGAATGACTCTTACAGAGAGAGAGTGTCCAGAG 780
QY 791 CTGGGTTTCAGAGATGTCTCACCGGACCAAGAGTGTCCCAAGAGATTTTGGCAACGC 850

```

```

Db 781 CTGGGTTTCAGAGATGTCTCACCGGACCAAGAGTGTCCCAAGAGATTTTGGCAACGC 840
QY 851 TTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCCCTCCACAGGTCTGAATGCCCT 910
Db 841 TTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCCCTCCACAGGTCTGAATGCCCT 900
QY 911 TCCAGCGGTATATCTCCCTCCAGTGTTCCTCACTGGCGAGCTGAGAGGCCATGACCGGCGG 970
Db 901 TCCAGCGGTATATCTCCCTCCAGTGTTCCTCACTGGCGAGCTGAGAGGCCATGACCGGCGG 960
QY 971 ATCGTGGAGGGGGGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTCACTTC 1030
Db 961 ATCGTGGAGGGGGGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTCACTTC 1020
QY 1031 GGCACCAACCCATCTGTGAGAGCAGCCTCATTTAGCCGCCAGTGGTCTCACTGCGGCC 1090
Db 1021 GGCACCAACCCATCTGTGAGAGCAGCCTCATTTAGCCGCCAGTGGTCTCACTGCGGCC 1080
QY 1091 CACTGCTTTCTTGTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGGTGTACCGCGGAC 1150
Db 1081 CACTGCTTTCTTGTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGGTGTACCGCGGAC 1140
QY 1151 AGCAACCTGACCAAGTGGCTGAGAGCAGCCTCCATTTGCCAGATCATCATCAAGCAAT 1210
Db 1141 AGCAACCTGACCAAGTGGCTGAGAGCAGCCTCCATTTGCCAGATCATCATCAAGCAAT 1200
QY 1211 TACACCGATGAGAGAGAGCATATGATGATGAGCCTCTTACGCGCTGTCCAGGCCCTGACC 1270
Db 1201 TACACCGATGAGAGAGAGCATATGATGATGATGAGCCTCTTACGCGCTGTCCAGGCCCTGACC 1260
QY 1271 CTGTGCGCTCATCATCACCTGCTTGCCTTCCATGCAATGAGACAGACCTTTAGCTCAAT 1330
Db 1261 CTGTGCGCTCATCATCACCTGCTTGCCTTCCATGCAATGAGACAGACCTTTAGCTCAAT 1320
QY 1331 GAGACCTGCTGATCAAGAGCTTTGGCAAGACAGGAGACAGATGACAAGATATCCGCC 1390
Db 1321 GAGACCTGCTGATCAAGAGCTTTGGCAAGACAGGAGACAGATGACAAGATATCCGCC 1380
QY 1391 TTCTCCGGAGAGGTGACAGTCAATCTGATGATCTTCAAGAAATGATGATCACTTGTGTC 1450
Db 1381 TTCTCCGGAGAGGTGACAGTCAATCTGATGATCTTCAAGAAATGATGATCACTTGTGTC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGATATGTGTCTGGGACCTTGTGGGGGACAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGATATGTGTCTGGGACCTTGTGGGGGACAGAGAC 1500
QY 1511 TCCGCGAAGGAGACAGCGGGGGGCTCTGTCTGTGAGAGACCAACCGCTGGTATCCTG 1570
Db 1501 TCCGCGAAGGAGACAGCGGGGGGCTCTGTCTGTGAGAGACCAACCGCTGGTATCCTG 1560
QY 1571 GCAAGTGTCAACAGCTGGGACAGAGCTGTGGCCAGAGAAACAAACCTGTGTGTACACC 1630
Db 1561 GCAAGTGTCAACAGCTGGGACAGAGCTGTGGCCAGAGAAACAAACCTGTGTGTACACC 1620
QY 1631 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAGAGAGTGTGATTCAG 1690
Db 1621 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAGAGAGTGTGATTCAG 1680
QY 1691 AAATCTTAA 1699
Db 1681 AAATCTTAA 1689

```

```

RESULT 6
CS081489
LOCUS CS081489 2380 bp DNA linear PAT 18-MAY-2005
DEFINITION Sequence 1 from Patent WO2005040401.
ACCESSION CS081489
VERSION CS081489.1 GI:66348810
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
Bayer Healthcare AG (DE)
Location/Qualifiers

1. .2380
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
Source

1. .2380

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 95.6%; Score 1670.4; DB 2; Length 2380;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CTCAGAGCCATGAGAGGAGACAGCCAGGGAATGATCTCCAGAGAAAGACCTTCAGC 60
78 CTCAGAGCCATGAGAGGAGACAGCCAGGGAATGATCTCCAGAGAAAGACCTTCAGC 137
61 TGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGC 120
138 TGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGC 197
121 CCAGGATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGC 180
198 CCAGGATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGC 257
181 CCAGGATCTCCAGCTGGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGC 240
258 CCAGGATCTCCAGCTGGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGC 317
241 CCAGGATCTCCAGGCGGGGATCTCCAGCTGGGACACCTTCAGGCTCCATCCAGC 300
318 CCAGGATCTCCAGGCGGGGATCTCCAGCTGGGACACCTTCAGGCTCCATCCAGC 377
301 CAGGTCATCATCCGCGAGGTCAAGCTCGGTGACAACTCCCAACAGAGTGTACTTGT 360
378 CAGGTCATCATCCGCGAGGTCAAGCTCGGTGACAACTCCCAACAGAGTGTACTTGT 437
361 TAGAGCAACACAGTGGGGGCTGTACCATCCGATCATCTCTCCAGAGTCAAGCACCAGC 420
438 TAGAGCAACACAGTGGGGGCTGTACCATCCGATCATCTCTCCAGAGTCAAGCACCAGC 497
421 AACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
498 AACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 557
481 CCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
558 CCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 617
541 CATCATCTCTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
618 CATCATCTCTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 677
601 GAGGTCATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
678 GAGGTCATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 737
661 CAGGTCATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
738 CAGGTCATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 797
721 CTCCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
798 CTCCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
781 CTCCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840

858 CTGCGACAGCTGGAGTTTCGAGAGTCTCACCGGACAAACGAGGTTGCCACAGGATTT 917
841 TGCCAAAGCTTTCATCTTTAGATACAACTCCACATCCAGGAAGCCCTCCACAGTTC 900
918 TGCCAAAGCTTTCATCTTTAGATACAACTCCACATCCAGGAAGCCCTCCACAGTTC 977
901 TGAATGCCCTTCCAGGAGTATATCTCTCCAGTGTTCCTCCAGGAGTGAAGGAGGAT 960
978 TGAATGCCCTTCCAGGAGTATATCTCTCCAGTGTTCCTCCAGGAGTGAAGGAGGAT 1037
961 GACCGGCGGATCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
1038 GACCGGCGGATCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1097
1021 TCTGACCTTGGGACCAACCACTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1098 TCTGACCTTGGGACCAACCACTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1157
1081 CACTGCGGCGGATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
1158 CACTGCGGCGGATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1217
1141 CGCGGAGACCAAGCACTGCAACGATTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
1218 CGCGGAGACCAAGCACTGCAACGATTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1277
1201 CAAAGCAATTAACCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
1278 CAAAGCAATTAACCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1337
1261 GCGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
1338 GCGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1397
1321 TAGCCTCAATGAGACCTGTGATGATCAAGGCTTTGGAGAGGAGGAGGAGGAGGAGG 1380
1398 TAGCCTCAATGAGACCTGTGATGATCAAGGCTTTGGAGAGGAGGAGGAGGAGGAGG 1457
1381 GACATCCGCTTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
1458 GACATCCGCTTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1517
1441 CTACTTGGTATGAGACCTGTGATGATCAAGGCTTTGGAGAGGAGGAGGAGGAGGAGG 1500
1518 CTACTTGGTATGAGACCTGTGATGATCAAGGCTTTGGAGAGGAGGAGGAGGAGGAGG 1577
1501 GGGCAGAGACTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
1578 GGGCAGAGACTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1637
1561 CTGCTACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
1638 CTGCTACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1697
1621 TGTGTACACCAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1672
1698 TGTGTACACCAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1749

RESULT 7
DD200135 2393 bp DNA linear PAT 19-JAN-2006
LOCUS
DEFINITION
METHODS and compositions for treating cardiovascular disease.
ACCESSION
DD200135.1 GI:85650814
VERSION
DD200135.1 JP 200516605-A/50.
KEYWORDS
JP 200516605-A/50.
SOURCE
unclassified
ORGANISM
unclassified
REFERENCE
1 (bases 1 to 2393)
AUTHORS
Acton,S.L., Healy,E., Sutaguriano,N., Galvin,K.M., Chun,M.,
Rodriguez,M.W., Donaghy,M., Perodin,J., and Logan,T.J.
TITLE
Methods and compositions for treating cardiovascular disease

JOURNAL

Patent: JP 2005516605-A 50 09-JUN-2005;
Millennium Pharmaceuticals Inc

COMMENT

OS Homosapien
PN JP 2005516605-A/50
PD 09-JUN-2005

PF 29-JAN-2003 JP 2003565410
PR 05-DEC-2002 US 60/431079, 03-JUN-2002 US 60/394130, PR
19-APR-2002 US 60/373861, 01-FEB-2002 US 60/353224, PR
05-DEC-2002 US 60/431042, 15-MAR-2002 US 60/364529, PR
04-SEP-2002 US 60/408070, 10-JUL-2002 US 60/394797, PR
21-AUG-2002 US 60/404904, 23-AUG-2002 US 60/405450, PR
06-NOV-2002 US 60/424300, 24-JUN-2002 US 60/390971, PR
12-JUN-2002 US 60/388080, 29-APR-2002 US 60/376287 PI Susan
I action, ellen healy, nancy sutaguriano, katherine m PI galvin,
PI myoung chun, amei rodorigyu-wei, marie domagyu, jacqueline PI
perodin,
PI thomas joseph logan

PI Key Location/Qualifiers.

FEATURES
source Location/Qualifiers

1..2393
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 95.6%; Score 1670.4; DB 2; Length 2393;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGAGAGCATGAGAGGAGGACGACGCGGAATGATCTTCAGACAGACACCTTCAGC 60
DB 78 CTCAGAGACCATGAGAGGAGGACGACGCGGAATGATCTTCAGACAGACACCTTCAGC 137
QY 61 TGGAGCATCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTCCAGCCCGGAGATCTCCAGC 120
DB 138 TGGAGCATCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTCCAGCCCGGAGATCTCCAGC 197
QY 121 CCAGGCACTCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTCCAGCCCGGAGATCTCCAGC 180
DB 198 CCAGGCACTCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTCCAGCCCGGAGATCTCCAGC 257
QY 181 CCAGGCACTCTCCAGCTGGATCACTCCAGCCCGGAGATCTCCAGCCCGGAGATCTCCAGC 240
DB 258 CCAGGCACTCTCCAGCTGGATCACTCCAGCCCGGAGATCTCCAGCCCGGAGATCTCCAGC 317
QY 241 CCAGGCACTCTCCAGCCCGGAGATCTCCAGCTGGAGACACCTCCAGCCCGGAGATCTCCAGC 300
DB 318 CCAGGCACTCTCCAGCCCGGAGATCTCCAGCTGGAGACACCTCCAGCCCGGAGATCTCCAGC 377
QY 301 CAGGTCATCATCCGCGAGGATCGGATCAAACTCCCAACAGAGTGTACTTGT 360
DB 378 CAGGTCATCATCCGCGAGGATCGGATCAAACTCCCAACAGAGTGTACTTGT 437
QY 361 TAGAGCAACACCAAGTGGGGCTGTACCCATCCGATCATCTCTCCAGGATCAAGAC 420
DB 438 TAGAGCAACACCAAGTGGGGCTGTACCCATCCGATCATCTCTCCAGGATCAAGAC 497
QY 421 AACAGGGGCCACAGGAGGAGCCAGGATCAAGGCTGCCAAGTTCACTCCGCGGAGG 480
DB 498 AACAGGGGCCACAGGAGGAGCCAGGATCAAGGCTGCCAAGTTCACTCCGCGGAGG 557
QY 481 CCAGAGCAGCTACCGCTCATCGGATGGATCTCTCTCTCAATTCCTCTGTGTTTGGCT 540
DB 558 CCAGAGAGGCTACCGCTCATCGGATGGATCTCTCTCTCAATTCCTCTGTGTTTGGCT 617
QY 541 CATCATCTCTTCCAGATTCTGGACAGGGCCACACAGGATCAAGGATCAAGAGACAGAGGA 600
DB 618 CATCATCTCTTCCAGATTCTGGACAGGGCCACACAGGATCAAGGATCAAGAGAGAGGA 677
QY 601 GAGCTGTCCCAAGACGCTGTTCCTGTGAGAGGGGTGGTGAATGCAAGCTGAAGAGTGA 660
DB 678 GAGCTGTCCCAAGACGCTGTTCCTGTGAGAGGGGTGGTGAATGCAAGCTGAAGAGTGA 737

QY 661 CGAGCTGGGCTGCGTGAAGTTTGAAGTGAAGGACCAAGTCTCTGCTTAAATCTACTCTGGGTC 720
DB 738 CGAGCTGGGCTGCGTGAAGTTTGAAGTGAAGGACCAAGTCTCTGCTTAAATCTACTCTGGGTC 797
QY 721 CTCCATCATGAGTGGCTTCCCATCTGTAGACCAACTGAAATGATCTCTTCAAGAGAAC 780
DB 798 CTCCATCATGAGTGGCTTCCCATCTGTAGACCAACTGAAATGATCTCTTCAAGAGAAC 857
QY 781 CTGCGACAGCTGGGTTTCCAGAGTGTCAACCGGACCAACCGAGGTTGCCACAGGATTT 840
DB 858 CTGCGACAGCTGGGTTTCCAGAGTGTCAACCGGACCAACCGAGGTTGCCACAGGATTT 917
QY 841 TGCCAAAGCTTCTCAATCTTGAATACAACTCCACCATCCAGGAAAGCTCTCAGAGTC 900
DB 918 TGCCAAAGCTTCTCAATCTTGAATACAACTCCACCATCCAGGAAAGCTCTCAGAGTC 977
QY 901 TGAATGCCCTTCCAGAGGATATATCTCCCTCAAGTGTCCCATCTGCGAATGAGGCCAT 960
DB 978 TGAATGCCCTTCCAGAGGATATATCTCCCTCAAGTGTCCCATCTGCGAATGAGGCCAT 1037
QY 961 GACCGGGCGAGTGGTGGAGAGGGGCGCTGGGCTCGGATAGCAAGTGGGCTTGGCAAGTGAG 1020
DB 1038 GACCGGGCGAGTGGTGGAGAGGGGCGCTGGGCTCGGATAGCAAGTGGGCTTGGCAAGTGAG 1097
QY 1021 TCTGCACTTGGGACCAACCAACATCTGTGAGAGGACGCTCATTTGACGCCAGTGGGATGCT 1080
DB 1098 TCTGCACTTGGGACCAACCAACATCTGTGAGAGGACGCTCATTTGACGCCAGTGGGATGCT 1157
QY 1081 CACTGCCGCCCACTGCTTCTTCTGTGACCCGGAGAAAGTCTTGAAGGCTGGAAAGTGTA 1140
DB 1158 CACTGCCGCCCACTGCTTCTTCTGTGACCCGGAGAAAGTCTTGAAGGCTGGAAAGTGTA 1217
QY 1141 CGGGGACACAGCAACCTGACACGATTGGCTGAGAGGACCTCCATTGCCAGATCATCAT 1200
DB 1218 CGGGGACACAGCAACCTGACACGATTGGCTGAGAGGACCTCCATTGCCAGATCATCAT 1277
QY 1201 CAACAGCAATTACACCGATGAGAGGAGCACTATGACATGCGCTCATGCGGCTGTCCAA 1260
DB 1278 CAACAGCAATTACACCGATGAGAGGAGCACTATGACATGCGCTCATGCGGCTGTCCAA 1337
QY 1261 GCCCTGACCTGTCCGCTCATATCCACCTGCTGCTGCCATGATGATGACAGACTT 1320
DB 1338 GCCCTGACCTGTCCGCTCATATCCACCTGCTGCTGCCATGATGATGACAGACTT 1397
QY 1321 TAGCCTCAATGAGACCTGTGATCAACAGGCTTGGCAAGACCAAGGATGACAA 1380
DB 1398 TAGCCTCAATGAGACCTGTGATCAACAGGCTTGGCAAGACCAAGGATGACAA 1457
QY 1381 GACATCCCTCTCTCCGAGAGGTGAGGATCAATCTGATGACTTCAAGAAATGCAATGA 1440
DB 1458 GACATCCCTCTCTCCGAGAGGTGAGGATCAATCTGATGACTTCAAGAAATGCAATGA 1517
QY 1441 CTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG 1500
DB 1518 CTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG 1577
QY 1501 GGGCAGAGACTCTCCGACAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAGAACCG 1560
DB 1578 GGGCAGAGACTCTCTGACAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAGAACCG 1637
QY 1561 CTGTTACTGGCAGAGTTCACCAAGCTGGGACACAGGCTGTGGCCAGAGAAACAACTGG 1620
DB 1638 CTGTTACTGGCAGAGTTCACCAAGCTGGGACACAGGCTGTGGCCAGAGAAACAACTGG 1697
QY 1621 TGTGTACACCAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGGAG 1672
DB 1698 TGTGTACACCAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGGAG 1749

RESULT 8
AB048796
LOCUS

AB048796

2393 bp

mRNA

linear

PRI 22-MAR-2001

```

DEFINITION Homo sapiens mspl mRNA for membrane-type mosaic serine protease,
complete cds.
ACCESSION AB048796
VERSION AB048796.1 GI:13429969
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 Kim,D.R., Sharmir,S., Inoue,M. and Kido,H.
TITLE Cloning and expression of novel mosaic serine proteases with and
without a transmembrane domain from human lung
JOURNAL Biochim. Biophys. Acta 1518 (1-2), 204-209 (2001)
PUBMED 11267681
REFERENCE 2 (bases 1 to 2393)
AUTHORS Kim,D.R., Inoue,M. and Kido,H.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2000) Hitoshi Kido, The University of Tokushima,
Division of Enzyme Chemistry, Institute for Enzyme Research,
Kutsumoto-cho, Tokushima 770, Japan
(E-mail:kido@eier.tokushima-u.ac.jp, Tel:81-88-633-7424,
Fax:81-88-633-7425)

FEATURES
source
    location/Qualifiers
        1..2393
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="lung"
            /clone_lib="Marathon Ready cDNA (Clontech)"
            1..2393
                /gene="mspl"
                1..87
                    /gene="mspl"
                    88..1833
                        /gene="mspl"
                        /codon_start=1
                        /transl_except=(pos:862..864,aa:Arg)
                        /transl_except=(pos:979..981,aa:His)
                        /transl_except=(pos:1573..1575,aa:His)
                        /product="membrane-type mosaic serine protease"
                        /protein_id="BAB39741.1"
                        /db_xref="GI:13429970"
                        /translation="MERDSHGNAAPARTPSAGASPAQASPATPPRGRASPQAASPAQAA
                        SPACTPRGRASPAQSPACPRGRASPRASPAQSPARASPALAISRSRGSSGSA
                        RASAVTTSPTRVTLVRAIPVAIVTRSGPASPARATPATRESPTSILPKRTREGQQQ
                        LPLGVCLLLIALVALSLILEFQWQGHTGLHKQRSCPCPHAVRCGVDDCKIKSDSE
                        LGCFRPMDKSLIKTVISGSSHQMPLICSSNMNDSYSEKTCQLGFESGAHRTEVAAHD
                        PANFSFLRYNSTIOESLHRSCHSORIYSQSHCGLRAMTGRIVGGALASDEKPMW
                        QVSLAFGTTHLCGGTLDIAQWLITLAHQFVLTRKEVLEGMVVVGATSNHLQLPKASII
                        AEIINSVTDEEDDYDIALMRLSKPILTISHIRPACLPMHGOTFSLNTKMTLTGPCK
                        TRFDIDKTSPLPIREYQNULDPFKCNDIPLYDSLITRRPMCAGIHGGRDSCQGSOGG
                        PLVEEQNNRWYLAVGWETGCGQRNKRPKVTTKEVLPMWTISKMEVRSLQDDTPSR
                        LGTSSGGPGGAPRV"
                1834..2393
                    /gene="mspl"
                    2343..2348
                        /gene="mspl"

```

| | | | | | |
|----|------|---------------------------|---|-------------------------------------|------|
| QY | 121 | CCAGGCATCTCCACACC | CCAGGCATCTCCAGCTGGGAA | CACTCCGGGCGGGGCATCTCCAGC | 180 |
| Db | 198 | CCAGGCATCTCCACACC | CCAGGCATCTCCAGCTGGGAA | CACTCCGGGCGGGGCATCTCCAGC | 257 |
| QY | 181 | CCAGGCATCTCCAGCTGGTA | CACCTCCAGGCGGGGACATCTCCAGGCGGGGCATCTCCAGC | 240 | |
| Db | 258 | CCAGGCATCTCCAGCTGGTA | CACCTCCAGGCGGGGACATCTCCAGGCGGGGCATCTCCAGC | 317 | |
| QY | 241 | CCAGGCATCTCCAGC | CCGGGCATCTCCGGCTCTGGGCATCACTTTCCAGGCTCTCAATCCGG | 300 | |
| Db | 318 | CCAGGCATCTCCAGC | CCGGGCATCTCCGGGCATCTCCGGGCATCTCCGGGCATCTTTCCAGGCTCTCAATCCGG | 377 | |
| QY | 301 | CAGGTCATCATCCGGC | CAGGTACGCTCGGTGACA | CACTCCCCAAACAGAGTGTACCTTGT | 360 |
| Db | 378 | CAGGTCATCATCCGGC | CAGGTACGCTCGGTGACA | CACTCCCCAAACAGAGTGTACCTTGT | 437 |
| QY | 361 | TAGAGCAACACAGTGGGGGCTGTAC | CCGATCCGATCTCTCGCAGGTAGACACAGC | 420 | |
| Db | 438 | TAGAGCAACACAGTGGGGGCTGTAC | CCGATCCGATCTCTCGCAGGTAGACACAGC | 497 | |
| QY | 421 | AACCAAGGCGCCAC | CAGGAGAGCCCAAGGTACAGACCTGGCCAAAGTTCA | CTTGGCGGAGGG | 480 |
| Db | 498 | AACCAAGGCGCCAC | CAGGAGAGAGCCCAAGGTACAGACCTGGCCAAAGTTCA | CTTGGCGGAGGG | 557 |
| QY | 481 | CCAGAGAGCTACCGCTCAT | CGGGTGGTGTCTCTCTCATTTGGCCTGTGTGTTCCT | 540 | |
| Db | 558 | CCAGAGAGCTACCGCTCAT | CGGGTGGTGTCTCTCTCTCATTTGGCCTGTGTGTTCCT | 617 | |
| QY | 541 | CATCATCTCTCTCCAGTTCT | GGCAGAGGCCACACAGGATCAAGGTACAGAGACAGAGGA | 600 | |
| Db | 618 | CATCATCTCTCTCCAGTTCT | GGCAGAGGGGCACACAGGATCAAGGACAAAGAGACAGAGGA | 677 | |
| QY | 601 | GAGCTGTCCCAAGAC | CGCTGTTCGCTGTGACGGGGTGGTGTGACTGCAAGCTGAAGATGA | 660 | |
| Db | 678 | GAGCTGTCCCAAGAC | CGCTGTTCGCTGTGACGGGGTGGTGTGACTGCAAGCTGAAGATGA | 737 | |
| QY | 661 | CGAGCTGGGCTGGCTG | AGATTGA | CTGGGACAAGTCTCTGTTAAATCTACTTGGGTC | 720 |
| Db | 738 | CGAGCTGGGCTGGCTG | AGATTGA | CTGGGACAAGTCTCTGTTAAATCTACTTGGGTC | 797 |
| QY | 721 | CTCCCATATAGTGGCTT | CCCATCTGTAGACAGAA | CTGGAATCTCTACTCAGAGAAAC | 780 |
| Db | 798 | CTCCCATATAGTGGCTT | CCCATCTGTAGACAGAA | CTGGAATCTCTACTCAGAGAAAC | 857 |
| QY | 781 | CTGCACAGCAGTGGGGTTT | CGAGTGTCTCA | CCGGAACAACAGGTTGGCCACAGGATTT | 840 |
| Db | 858 | CTGCACAGCAGTGGGGTTT | CGAGTGTCTCA | CCGGAACAACAGGTTGGCCACAGGATTT | 917 |
| QY | 841 | TGCCAACAGCTTCTCAATCTT | GAGATAC | CACTCCACATCTCCAGGAAAGCCTTCCACAGGTC | 900 |
| Db | 918 | TGCCAACAGCTTCTCTCAATCTT | GAGATAC | CACTCCACATCTCCAGGAAAGCCTTCCACAGGTC | 977 |
| QY | 901 | TGAATGCCCTTCCAGCGGGTA | TATCTCTCCAGTGTTC | CCCATCTGCGGACTGAGGGGCAT | 960 |
| Db | 978 | TGAATGCCCTTCCAGCGGGTA | TATCTCTCTCCAGTGTTC | CCCATCTGCGGACTGAGGGGCAT | 1037 |
| QY | 961 | GACCGGGCGGATTCG | TGGGAGGGGGGCTGGGCTCGGATG | CAAGGAGGCTTGGCAAGTAG | 1022 |
| Db | 1038 | GACCGGGCGGATTCG | TGGGAGGGGGGCTGGGCTCGGATG | CAAGGAGGCTTGGCAAGTAG | 1097 |
| QY | 1021 | TCTGCACCTTCGGGAC | CAACCCACATCTGTGAGGAC | CGCTCAATTGACGCCCAAGTGGTGTCT | 1080 |
| Db | 1098 | TCTGCACCTTCGGGAC | CAACCCACATCTGTGAGGAC | CGCTCAATTGACGCCCAAGTGGTGTCT | 1155 |
| QY | 1081 | CACGTCCGCCCACTGTCTT | TCGTGAC | CCCGGGAAGAAGTCTTGGAGGGCTTGAAGAAGTGT | 1140 |
| Db | 1158 | CACGTCCGCCCACTGTCTT | TCGTGAC | CCCGGGAAGAAGTCTTGGAGGGCTTGAAGAAGTGT | 1217 |
| QY | 1141 | CGCGGGACCAAGCAAC | CTGACACAGTTGCC | TGAGGACGCTTCCATTTGCCGAGATCATAT | 1200 |
| Db | 1218 | CGCGGGACCAAGCAAC | CTGACACAGTTGCC | TGAGGACGCTTCCATTTGCCGAGATCATAT | 1277 |
| QY | 1201 | CAACGACATTAC | CCGATGAGAGAGAC | GACTATGACATTCGCTCATTCGCGCTGTCTCA | 1266 |

| | | | |
|------------|---|--|-----------------------------|
| Db | 1278 | CACAGCAATTACACCATGAGAGAGACACTAATGACTCGCCCTCATAGCGCTGTCCA | 1337 |
| Qy | 1261 | GCCCCTAGCCTGTCCGCTCAATCCACCTTCCTGCCATGATGACAGACCTT | 1320 |
| Db | 1338 | GCCCCTAGCCTGTCCGCTCAATCCACCTTCCTGCCATGATGACAGACCTT | 1397 |
| Qy | 1321 | TAGCCTCAATGAGACCTGTGGATTCACAGCCTTTGGCAAAGCACAGGAGACAGATGACAA | 1386 |
| Db | 1398 | TAGCCTCAATGAGACCTGTGGATTCACAGCCTTTGGCAAAGCACAGGAGACAGATGACAA | 1457 |
| Qy | 1381 | GACATCCCCTCTTCCTCCGGAAGTGTCAGATCAATCTCATGCATCTTCAAGAAATGCATGA | 1448 |
| Db | 1458 | GACATCCCCTCTTCCTCCGGAAGTGTCAGATCAATCTCATGCATCTTCAAGAAATGCATGA | 1517 |
| Qy | 1441 | CTACTTGGTCTATGACAGTTACCTTAACCCCAAGATGATGTGTGCTGGGACCTTCGTGG | 1500 |
| Db | 1518 | CTACTTGGTCTATGACAGTTACCTTAACCCCAAGATGATGTGTGCTGGGACCTTCGTGG | 1577 |
| Qy | 1501 | GGGACAGAGACTCTGCGCCAGAGGAGACAGCGGGGGGGCTCTTGTCTGTGAGCAGAACAACG | 1566 |
| Db | 1578 | GGGACAGAGACTCTCTCCAGAGGAGACAGCGGGGGGGCTCTTGTCTGTGAGCAGAACAACG | 1637 |
| Qy | 1561 | CTGTACTCTGCAAGGTGTCAACCACTGCGGCAACAGCTGTGTGGCCAGAGAAAACAACTGG | 1620 |
| Db | 1638 | CTGTACTCTGCAAGGTGTCAACCACTGCGGCAACAGCTGTGTGGCCAGAGAAAACAACTGG | 1697 |
| Qy | 1621 | TGTGTACCAACCAAGTACAGAAAGTTCTTCCCTGGAATTACAGCAAGATGGAG | 1672 |
| Db | 1698 | TGTGTACCAACCAAGTACAGAAAGTTCTTCCCTGGAATTACAGCAAGATGGAG | 1749 |
| RESULT 9 | | | |
| LOCUS | AY190317 | 2237 bp | mRNA linear PRI 12-FEB-2003 |
| DEFINITION | Homo sapiens transmembrane protease serine 6 (TMPRSS6) mRNA, | | |
| ACCESSION | AY190317 | | |
| VERSION | AY190317.1 | | |
| KEYWORDS | GI:28371867 | | |
| SOURCE | | | |
| ORGANISM | Homo sapiens (human) | | |
| | Bukariya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. | | |
| REFERENCE | 1 (bases 1 to 2237) Park,T.J. and Park,W.J. Homo sapiens transmembrane protease, serine 6 (TMPRSS6) mRNA Unpublished | | |
| AUTHORS | 2 (bases 1 to 2237) Park,T.J. and Park,W.J. Direct Submission Submitted (02-DEC-2002) Life Science, K-JIST, Oryoung-dong 1, Kwangju 500-712, Rep. Of Korea | | |
| JOURNAL | location/Qualifiers | | |
| FEATURES | | | |
| SOURCE | 1..2237 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="mRNA" | | |
| | /db_xref="taxon:9606" | | |
| | /chromosome="11" | | |
| | /map="11q23" | | |
| | 1..2237 | | |
| | /gene="TMPRSS6" | | |
| | 89..1765 | | |
| | /gene="TMPRSS6" | | |
| | /codon_start=1 | | |
| | /product="transmembrane protease serine 6" | | |
| | /protein_id="AA038062.1" | | |
| | /translation="MERDSSHGNASPARTPSAQSPQAQSPAGTTPGRASPQAQSPAQAQ SPAGTPPGRASPQRASPPRASAQSPARASPALASLSRSSSGSSSS RSASVTSPTVRVAVRAVPVAIVIRSPASAPATRRESPTGISLPKTRMGOKK LPLRGVLLALTVALLIILFQFMQGYIGYIKERESCPHAVARIDVDCKSKDEB | | |
| gene | | | |
| CDS | | | |

[illegible]


```

Db      799  CTCCCATAGTGGCTCCCATCTGTAGAGCACTGAAATGACTCTACTCAGAAAGAC 858
Qy      781  CTGACAGAGCTGGGTTTCCAGAGTGTCTACCCGAGCAACGAGTTGGCCACAGGATTT 840
Db      859  CTGACAGAGCTGGGTTTCCAGAGTGTCTACCCGAGCAACGAGTTGGCCACAGGATTT 918
Qy      841  TGCCAAAGCTTCTCAATCTTGAATACCACTCCACCATCCAGAAAGCTCCACAGTTC 900
Db      919  TGCCAAAGCTTCTCAATCTTGAATACCACTCCACCATCCAGAAAGCTCCACAGTTC 978
Qy      901  TGAATGCCCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCCAGTCCAGTCCAGTCC 960
Db      979  TGAATGCCCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCCAGTCCAGTCCAGTCC 1038
Qy      961  GACCGGGGCGATCTGTGGAGGGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAATGAG 1020
Db      1039  GACCGGGGCGATCTGTGGAGGGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTAG 1098
Qy      1021  TCTGACCTTGGGAGCAACCCCATCTGTGAGGACGCTCATTTGACGCCAGTGGTGTCT 1080
Db      1099  TCTGACCTTGGGAGCAACCCCATCTGTGAGGACGCTCATTTGACGCCAGTGGTGTCT 1158
Qy      1081  CACTGCCGCCACTGTCTTCTGTGACCCGGGAGAAAGTCTCTGAGGGGCTGGAAGTGT 1140
Db      1159  CACTGCCGCCACTGTCTTCTGTGACCCGGGAGAAAGTCTCTGAGGGGCTGGAAGTGT 1218
Qy      1141  CGCGGGGACCAAGCACTGTGACCGAGTTGCTTGGAGGACGCTTCCATTGGCCGATCAT 1200
Db      1219  CGCGGGGACCAAGCACTGTGACCGAGTTGCTTGGAGGACGCTTCCATTGGCCGATCAT 1278
Qy      1201  CAAAGCAATTAACCAAGTGAAGGAGAGCAATATAGACATGCGCCCTCATGCGGCTGTCAA 1260
Db      1279  CAAAGCAATTAACCAAGTGAAGGAGAGCAATATAGACATGCGCCCTCATGCGGCTGTCAA 1338
Qy      1261  GCCCCTGACCTGTCCGCTCATCTCAACCCCTGTGCTCCCTCCATGATGACAGACTT 1320
Db      1339  GCCCCTGACCTGTCCGCTCATCTCAACCCCTGTGCTCCCTCCATGATGACAGACTT 1398
Qy      1321  TAGCCTCAATGAGACTGTGTGATCAGAGCTTTGGCAAGCAAGGAGAGACAGATGACAA 1380
Db      1399  TAGCCTCAATGAGACTGTGTGATCAGAGCTTTGGCAAGCAAGGAGAGACAGATGACAA 1458
Qy      1381  GACATCCCCCTCTCTCCGGGAGGAGTGAATCTCATGCACTTCAAGAAATGCATGTA 1440
Db      1459  GACATCCCCCTCTCTCCGGGAGGAGTGAATCTCATGCACTTCAAGAAATGCATGTA 1518
Qy      1441  CTACTGGTGTATGAGAGTACCTTACCCCAAGATGATGTGTCTGGGAGCTTCTGTGG 1500
Db      1519  CTACTGGTGTATGAGAGTACCTTACCCCAAGATGATGTGTCTGGGAGCTTCTGTGG 1578
Qy      1501  GGGGAGAGACTCTGTGCAAGGAGAGCAAGCGGGGGCTTGTGTGTGAGCAAGCAACCG 1560
Db      1579  GGGGAGAGACTCTGTGCAAGGAGAGCAAGCGGGGGCTTGTGTGTGAGCAAGCAACCG 1638
Qy      1561  CTGTGTACTGTGAGAGTGTCAACAGTGGGGAGACAGCTGTGGCCAGAGAAACAACCTGG 1620
Db      1639  CTGTGTACTGTGAGAGTGTCAACAGTGGGGAGACAGCTGTGGCCAGAGAAACAACCTGG 1698
Qy      1621  TGTGTACACCAAAATGACAGAAATTTCTTCCCTGATTTTACAGAAAGATGAGAGAGT 1680
Db      1699  TGTGTACACCAAAATGACAGAAATTTCTTCCCTGATTTTACAGAAAGATGAGAGAGT 1758
Qy      1681  GCGATTGAGAA 1691
Db      1759  GGGTTGAAAAA 1769

```

```

RESULT 10
CS075313  CS075313  2067 bp  DNA  linear  PAT 05-MAY-2005
LOCUS      Sequence 1 from Patent WO2005019270.
DEFINITION CS075313
ACCESSION  CS075313.1  GI:63091664
VERSION

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

```

1
Madison, E. L. and Nixon, A.
Endothelase-2 ligands
Patent: WO 2005019270-A 1 03-MAR-2005,
Dyax Corp. (US); Dendreon Corporation (US)
Location/Qualifiers
1..2067
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..2067
/note="unnamed protein product"
/codon_start=1
/protein_id="CA193712.1"
/db_xref="GI:63091665"
/translation="MERDSHGASPARTSPASPAQASPARASPALASLSRSSGSSA
SPAGTPGRASPAQASPAQTPPGRASPAQASPARASPALASLSRSSGSSA
RSASVTPSPRVAVRATPVGAVPIRSSPARSAPATATSPESPTSLPKFTWREGQO
LPLIGCVLLIALVSLIILFOPMOGHTGIRYKORESCPHAVRCGVDDKLSDE
LGVVRPDMDSILIKIYSSSHOWLPICSSMWNDSSEKTCOOLGFSFAHRTTEVAHD
FANSFSLRYNSTIOESLHRSRCPORSYISLQSHGCLRAMTRIRIVGALASDKPW
QVSLHFGTHICGTLIDQWVLTAAHCFVTEKEVLEKVKVYAGTSNHLQLEPAASI
AETIISNNTDEDDVDIALMRLSKPLTSLAH,HPACLPMHGOFTSLNCTWTGFGK
TRETDDKTSPLREVQVNLIDPFKCNLDLYDSVLELPRMNCAGDLRGDRSCODSG
PLVEONLRMYLAVGTSWGTGCGORNRPGVYTVTELELPWYISMEKRAORVEKAWY
RPRQLGRMSYPSIFLCRYAMPENVSVAEDPVYFVYKHLCMGRSSMPPLAV
LVPLIFPLLLSLFKNTSDSILTLTFTVATMLPENYHSPFPLKIFHONLTITIK
EYRVIK"

ORIGIN

Query Match 95.2%; Score 1663.4; DB 2; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      11  ATGAGAGGAGACAGCCAGGGAATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 70
Db      1  ATGAGAGGAGACAGCCAGGGAATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 60
Qy      71  CCAAGCCAGCATCTCCAGCTGGAGACACTTCAGAGCCGGGCAATCTCCAGCCAGGATCT 130
Db      61  CCAAGCCAGCATCTCCAGCTGGAGACACTTCAGAGCCGGGCAATCTCCAGCCAGGATCT 120
Qy      131  CCAAGCCAGCATCTCCAGCTGGAGACACTTCAGAGCCGGGCAATCTCCAGCCAGGATCT 190
Db      121  CCAAGCCAGCATCTCCAGCTGGAGACACTTCAGAGCCGGGCAATCTCCAGCCAGGATCT 180
Qy      191  CCAAGCTGTACACTCCAGGACCGGGCAATCTCCAGAGCCGGGCAATCTCCAGCCAGGATCT 250
Db      181  CCAAGCTGTACACTCCAGGACCGGGCAATCTCCAGAGCCGGGCAATCTCCAGCCAGGATCT 240
Qy      251  CCAAGCCGGGCAATCTCCAGCTGGAGACACTTCAGAGCTTCATCTCCAGGAGTATCA 310
Db      241  CCAAGCCGGGCAATCTCCAGCTGGAGACACTTCAGAGCTTCATCTCCAGGAGTATCA 300
Qy      311  TCCGCAAGTCAAGCTCGGAGCAACTCCCAACAGAGTATCTTGTATAGACAACA 370
Db      301  TCCGCAAGTCAAGCTCGGAGCAACTCCCAACAGAGTATCTTGTATAGACAACA 360
Qy      371  CCAAGTGGGGCTGTACCATCCGATCATCTCTGCGCAGGTCAAGCAGCAGCAACCAAGGGCC 430
Db      361  CCAAGTGGGGCTGTACCATCCGATCATCTCTGCGCAGGTCAAGCAGCAGCAACCAAGGGCC 420
Qy      431  ACCAAGGAGAGCCAGGTATGAGAGCTGCCCAAGTTCACTTGGGGAGGGCCAGAAACAG 490
Db      421  ACCAAGGAGAGCCAGGTATGAGAGCTGCCCAAGTTCACTTGGGGAGGGCCAGAAACAG 480
Qy      491  CTACCGCTCATCGGAGTGTCTCTCTCATATGCGCTGTGTGTTTGCTCATCATCTCTC 550

```

```

Db      481  CTACCGCTCATCGGGTGGTCTCTCTCTCATTCGCTGGTGTTCCTCATCTCTC 540
Qy      551  TTCCAGTTCTGGACGGCCACACAGAGATCATGATCAAGAGACAGAGAGAGCTGCC 610
Db      541  TTCCAGTTCTGGACGGCCACACAGAGATCATGATCAAGAGACAGAGAGAGCTGCC 600
Qy      611  AAGACGCTGTTCCTGTGTACCGGGGTGGTGAAGTCAAGCTGAAGAGTGAAGCTGG 670
Db      601  AAGACGCTGTTCCTGTGTACCGGGGTGGTGAAGTCAAGAGTGAAGAGTGAAGCTGG 660
Qy      671  TGCGTGAAGTTTGAAGTGAAGAGTCTGTCTTAAATCTACTCTGGGTCTCTCCATCAG 730
Db      661  TGCGTGAAGTTTGAAGTGAAGAGTCTGTCTTAAATCTACTCTGGGTCTCTCCATCAG 720
Qy      731  TGCGTTCCTCATCTGTAGACAGCACTGGAATGCTCTCTAGAGAAAGCTGCGACAG 790
Db      721  TGCGTTCCTCATCTGTAGACAGCACTGGAATGCTCTCTAGAGAAAGCTGCGACAG 780
Qy      791  CTGGGTTTGAAGAGTCTACCGGACAAACGAGGTTGCCCAAGGATTTTGGCAACAGC 850
Db      781  CTGGGTTTGAAGAGTCTACCGGACAAACGAGGTTGCCCAAGGATTTTGGCAACAGC 840
Qy      851  TTCTCAATCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 910
Db      841  TTCTCAATCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 900
Qy      911  TCCGACGCGTATATCTCCCTCAGTGTCCCACTGCGAGTGAAGGACCATGCGGCG 970
Db      901  TCCGACGCGTATATCTCCCTCAGTGTCCCACTGCGAGTGAAGGACCATGCGGCG 960
Qy      971  ATCGTGAAGGGGCGCTGCGCTCGATGAAAGTGGCTTGGCAAGTGTCTGCACTTC 1030
Db      961  ATCGTGAAGGGGCGCTGCGCTCGATGAAAGTGGCTTGGCAAGTGTCTGCACTTC 1020
Qy      1031  GGACACACCATCTGTGAGAGGACGCTCATTAAGCGCCAGTGGGTCTCACTGCGCC 1090
Db      1021  GGACACACCATCTGTGAGAGGACGCTCATTAAGCGCCAGTGGGTCTCACTGCGCC 1080
Qy      1091  CACTGCTTCTTCGTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGTGTACGCGG 1150
Db      1081  CACTGCTTCTTCGTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGTGTACGCGG 1140
Qy      1151  AGCAACCTGACCAAGTTGCTGAGGACGCTTCATTCGCAAGTCAATCATCAAGCAAT 1210
Db      1141  AGCAACCTGACCAAGTTGCTGAGGACGCTTCATTCGCAAGTCAATCATCAAGCAAT 1200
Qy      1211  TACACCGATGAGAGGACGATATGACATGCGCCCTCAAGCGGCTGTCCAAAGCTGACC 1270
Db      1201  TACACCGATGAGAGGACGATATGACATGCGCCCTCAAGCGGCTGTCCAAAGCTGACC 1260
Qy      1271  CTGTCCGCTCATCATCAACCTGCTTGCCTCCCATGATGACAGACCTTTAGCTCAAT 1330
Db      1261  CTGTCCGCTCATCATCAACCTGCTTGCCTCCCATGATGACAGACCTTTAGCTCAAT 1320
Qy      1331  GAGACCTGCTGATCAAGGCTTTTGGCAAGACAGAGAGACAGATGACAAAGATCTCCC 1390
Db      1321  GAGACCTGCTGATCAAGGCTTTTGGCAAGACAGAGAGACAGATGACAAAGATCTCCC 1380
Qy      1391  TTCTCCGAGGAGTGAAGTGAATCTCATGCACTTCAAGAAATGCAATGATTAATGCTC 1450
Db      1381  TTCTCCGAGGAGTGAAGTGAATCTCATGCACTTCAAGAAATGCAATGATTAATGCTC 1440
Qy      1451  TATGACAGTTACTTACCCCAAGATGATGTGTGTGGGAGCTTTCGTGGGGGAGAGAC 1510
Db      1441  TATGACAGTTACTTACCCCAAGATGATGTGTGTGGGAGCTTTCGTGGGGGAGAGAC 1500
Qy      1511  TCCTGACAGGAGACAGCGGGGGGCTCTTGTGTGTGAGCAAGCAACCGTGTGTAATG 1570
Db      1501  TCCTGACAGGAGACAGCGGGGGGCTCTTGTGTGTGAGCAAGCAACCGTGTGTAATG 1560
Qy      1571  GCAGGTGTACACACTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACAC 1630

```

```

Db      1561  GCAGGTGTACACACTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACAC 1630
Qy      1631  AAGTGAAGAGTCTTCCCTGATTTACAGCAAGATGAGAGC 1675
Db      1621  AAGTGAAGAGTCTTCCCTGATTTACAGCAAGATGAGAGC 1665

RESULT 11
DD194636
LOCUS      DD194636
DEFINITION Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
ACCESSION DD194636
VERSION   DD194636.1 GI:85634727
KEYWORDS  JP 2005518332-A/13.
SOURCE    unidentified
ORGANISM  unidentified
REFERENCE 1 (bases 1 to 2067)
AUTHORS   Komandla,M., Sempke,J.E., Vlaeuk,G.P., Kemp,S.J., Slev,D.V. and
          Madison,B.L.
TITLE     Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
          Thereof
JOURNAL   Patent: JP 2005518332-A 13 23-JUN-2005:
          Edwin L. Madison, Joseph Edward Sempke, George P Vlaeuk, Scott Jeffrey
          Kemp, Mallareddy Komandla, Daniel Vanna Slev
          OS Homo Sapien
          PN JP 2005518332-A/13
          PD 23-JUN-2005
          PF 23-MAY-2002 JP 2002592470
          PR 23-MAY-2001 US 60/293267
          PI mallareddy komandla, joseph edward sempke, george p vlaeuk, pi
          scott jeffrey kemp, daniel vanna slev, edwin l madison cc
          FH Key Location/Qualifiers
          FT CDS (1) (2067)
          /note="Nucleic acid encoding (endorhelase FT
          2-L) protein".
FEATURES
source location/Qualifiers
1..2067
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 95.2%; Score 1663.4; DB 2; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

11 ATGAGAGGGACAGCCACGGGATGCAATCTTCACAGCAAGAACACTTCAAGTGAAGCATCT 70
Db 1 ATGAGAGGGACAGCCACGGGATGCAATCTTCACAGCAAGAACACTTCAAGTGAAGCATCT 60
Qy 71 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGCCAGGACATCT 130
Db 61 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGCCAGGACATCT 120
Qy 131 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGCCAGGACATCT 190
Db 121 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGCCAGGACATCT 180
Qy 191 CCAGCTGTACACTCCAGGCGGGGATCTTCAGAGCGGGGATCTTCAGAGCCAGGACATCT 250
Db 181 CCAGCTGTACACTCCAGGCGGGGATCTTCAGAGCGGGGATCTTCAGAGCCAGGACATCT 240
Qy 251 CCAGCCGGGATCTCCAGGCTTGGGATCACTTTCAGAGTCTTCATCCGGGAGGATCA 310
Db 241 CCAGCCGGGATCTCCAGGCTTGGGATCACTTTCAGAGTCTTCATCCGGGAGGATCA 300
Qy 311 TCCGCAAGTCAAGCTCGGTGACAACTCCCAACAGAGTGTACTTGTGTAAGACA 370
Db 301 TCCGCAAGTCAAGCTCGGTGACAACTCCCAACAGAGTGTACTTGTGTAAGACA 360
Qy 371 CCAAGTGGGGCTGTACCAATCCGATCATCTCTGCCAGGTGACGACCAAGCAAGGAGCC 430

```

| | | | |
|----|------|--|------|
| Dp | 361 | CCAGTGGGGGCTGTATCCATCCGATCATCTCTGCCAGGTAGGACACGACACAGGGCC | 420 |
| Qy | 431 | ACCAAGGAGAGACCCAGGTACGAGCCCTGCCAAGTTCACTTGCGGGAGAGGCGCAGAAGCAG | 490 |
| Dp | 421 | ACCAAGGAGAGACCCAGGTACGAGCCCTGCCAAGTTCACTTGCGGGAGAGGCGCAGAAGCAG | 480 |
| Qy | 491 | CTACCGCTCATTCGGGTGCGTGTCTCTCTCATTTGCTCGTGTGTTTCGCTCATCATCTCTC | 550 |
| Dp | 481 | CTACCGCTCATTCGGGTGCGTGTCTCTCTCATTTGCTCGTGTGTTTCGCTCATCATCTCTC | 540 |
| Qy | 551 | TTCCAGTTCGTGGACGGGCGACACAGGGATCGGTACAAAGAGACAGAGGAGAGGCTGTCC | 610 |
| Dp | 541 | TTCCAGTTCGTGGACGGGCGACACAGGGATCGGTACAAAGAGACAGAGGAGAGGCTGTCC | 600 |
| Qy | 611 | AAGCAGCGTGTTCGCTGTGACGGGGTGTGTGACTGCAAGCTGAAGAGTGAAGAGCTGTGGC | 670 |
| Dp | 601 | AAGCAGCGTGTTCGCTGTGACGGGGTGTGTGACTGCAAGCTGAAGAGTGAAGAGCTGTGGC | 660 |
| Qy | 671 | TGCGTAGAGTTTGAACGTGGACAAAGCTCTGCTTAAATCTACTGTGGTCTCTCCCATCAG | 730 |
| Dp | 661 | TGCGTAGAGTTTGAACGTGGACAAAGCTCTGCTTAAATCTACTGTGGTCTCTCCCATCAG | 720 |
| Qy | 731 | TGCGTTCCTCATCTGTAGAGCAACTGGAAATGACTCTACTAGAGAAAGCTGCCAGCAG | 790 |
| Dp | 721 | TGCGTTCCTCATCTGTAGAGCAACTGGAAATGACTCTACTAGAGAAAGCTGCCAGCAG | 780 |
| Qy | 791 | CTGGGTTTCGAGAGTGCCTACCGGACAAACGAGGTTGCCACAGGGATTTTGGCAACAGC | 850 |
| Dp | 781 | CTGGGTTTCGAGAGTGCCTACCGGACAAACGAGGTTGCCACAGGGATTTTGGCAACAGC | 840 |
| Qy | 851 | TTCTCAATCTTGAGATACAACTCCACCACTCCAGGAAAGCTTCCAAGGCTGAAATGCCCT | 910 |
| Dp | 841 | TTCTCAATCTTGAGATACAACTCCACCACTCCAGGAAAGCTTCCAAGGCTGAAATGCCCT | 900 |
| Qy | 911 | TCCCAAGCGGTATATCTCCCTCCAGTGTTCCTCACTGCGGACTGAGGGCCATGACCGGGCGG | 970 |
| Dp | 901 | TCCCAAGCGGTATATCTCCCTCCAGTGTTCCTCACTGCGGACTGAGGGCCATACCGGGCGG | 960 |
| Qy | 971 | ATCGGGGAGGGGGCGCTGGCTCCGGATGAGAAAGGAGCTTGGCAAGTGAAGCTGCACCTTC | 1030 |
| Dp | 961 | ATCGGGGAGGGGGCGCTGGCTCCGGATGAGAAAGGAGCTTGGCAAGTGAAGCTGCACCTTC | 1020 |
| Qy | 1031 | GGCACCACCCACATCTGTGAGAGGACGCTCATTTGACCGCCAGGTGGTTCCTCATGCGGCC | 1090 |
| Dp | 1021 | GGCACCACCCACATCTGTGAGAGGACGCTCATTTGACCGCCAGGTGGTTCCTCATGCGGCC | 1080 |
| Qy | 1091 | CATGCTCTTTCTGTGACCCCGGAGAGAGGTCTTGGAGGGCTGGAAAGGTATACGGGGCACCC | 1150 |
| Dp | 1081 | CATGCTCTTTCTGTGACCCCGGAGAGAGGTCTTGGAGGGCTGGAAAGGTATACGGGGCACCC | 1140 |
| Qy | 1151 | AGCAACCTGCACAGAGTTGGCTTGAGGACAGCTCCATTTGCCAGATCATATCAACAGCAAT | 1210 |
| Dp | 1141 | AGCAACCTGCACAGAGTTGGCTTGAGGACAGCTCCATTTGCCAGATCATATCAACAGCAAT | 1200 |
| Qy | 1211 | TACACCGATGAGAGAGAGCATATATCAATTCGCCCTCATGCGGCTGTCCAAAGCCCTGAC | 1270 |
| Dp | 1201 | TACACCGATGAGAGAGAGCATATATCAATTCGCCCTCATGCGGCTGTCCAAAGCCCTGAC | 1260 |
| Qy | 1271 | CTGTCCGCTCAATCCACCCCTGTGCTTCCCTCCCATGACATGACAGACCTTTAGCTCATAT | 1330 |
| Dp | 1261 | CTGTCCGCTCAATCCACCCCTGTGCTTCCCTCCCATGACATGACAGACCTTTAGCTCATAT | 1320 |
| Qy | 1331 | GAGACCTGCTGATATACAGGCTTTTGGCAAGACAGGGAGACAGATGACAAACATCCCCC | 1390 |
| Dp | 1321 | GAGACCTGCTGATATACAGGCTTTTGGCAAGACAGGGAGACAGATGACAAACATCCCCC | 1380 |
| Qy | 1391 | TTCTCTCCGGAGGTGCAAGTCAATCTCATCACTTCAAGAAATGCAATGACTACTTGGTTC | 1450 |
| Dp | 1381 | TTCTCTCCGGAGGTGCAAGTCAATCTCATCACTTCAAGAAATGCAATGACTACTTGGTTC | 1440 |
| Qy | 1451 | TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGACCTTTCGTGGGGCGAGAGAC | 1510 |

| | | | |
|----------------------------|--|--|----------------------------|
| Db | 1441 | TATACAGATTACCTTACCOCCAAGATGATGTGTGTCGGGACCTTGTTGGGGCAGAGAC | 1500 |
| Oy | 1511 | TCCTGCCAAGGGAGACAGCGGGGGCCCTCTTGTCTGTGAGCAACAACCCCTGGTACTCTG | 1570 |
| Db | 1501 | TCCTGCCAAGGGAGACACAGCGGGGGCCCTCTTGTCTGTGAGCAACAACCCCTGGTACTCTG | 1560 |
| Oy | 1511 | GCAGGTGTCAACCACTBTGGGGCACAGCGCTGTGGCCACAGAACAACAACTGGTGTACACC | 1630 |
| Db | 1561 | GCAAGTGTCAACCACTBTGGGGCACAGCGCTGTGGCCACAGAACAACAACTGGTGTACACC | 1620 |
| Oy | 1631 | AAAGTGACAGAAAGTTCTTCCCTGGATTTAAGACAGATGAGAGC | 1675 |
| Db | 1621 | AAAGTGACAGAAAGTTCTTCCCTGGATTTAAGACAGAAATGAGAAC | 1665 |
| <hr/> | | | |
| RESULT 12 | | | |
| LOCUS | AXI49581 | 2067 bp | DNA linear PAT 31-AUG-2001 |
| DEFINITION | Sequence 5 from Patent WO0136604. | | |
| ACCESSION | AXI49581 | | |
| VERSION | AXI49581.1 | GI:14348015 | |
| KEYWORDS | . | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominae; Homo. | | |
| REFERENCE | 1. Madison, E.L. and Ong, E.O. Nucleic acids encoding endothelialases, endothelialases and uses thereof Patent: WO 0136604-A 5 25-MAY-2001; | | |
| AUTHORS | CORVAS INTERNATIONAL, INC. (US) | | |
| JOURNAL | location/Qualifiers | | |
| FEATURES | 1..2067 | | |
| source | /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" | | |
| CDS | 1..2067 /note="unnamed protein product; DNA sequence encoding a transmembrane serine protease (endothelase-L) protein" /codon_start=1 /protein_id="CAC1221.1" /db_xref="GI:14348016" /translation="MERDSDSGNSPARTPSPASPAQSPAGTPPGSRASPAPQA SPAGTPGRASPAPQSPAGTPPGSRASPAPQSPAPASPLASLSRSGSSSA RASAVTSPTRVLYLVAITPGAVPIESPAPASA PAPTAPESGTSILPKRTMGCKO LPLGCFWDLIALVALVSLILIPFMQGHTGIRYEQRESCKRAVRCDGVDCUKSE LGCRFPWDKSLKLTYSGSOWLPICSSWNWDSYSEKTOQLGFESAHRTTEVAHD FANSFLTRYNSTIOBSLHRSCEPSQHYILQCSHCGLRMTRIVGALASDKPW QVSHFGFTTHICGGTLIDAOMWLVTAACHFVETREKVLEGKWAGTSLNHQLDEASI ARIITNSNTDEEDVDIALMRLSKPLTSLASHHPACLPMHGOTFSLNETCWITGEK TREIDDTSPPLEBEVDONLIDPFKCNDYLVDYISLPRMNCADLRGRSPCOSDSG PLVECONNRNYLAGVTWSWGTCGQRNPVGYYTVTEVLPIWYISKEMRARVERAMTY RVFGQLGRCSPRSIFLCXKAMPENVSVAEDFVIFYVKHLCMGIRSSWPALFY LVFFIFLLILLSLFKNTSDSIITLTFTFAVTMLPENVSFPPLKIIFQNLTIIK EYKIK" | | |
| <hr/> | | | |
| ORIGIN | | | |
| Query Match | 95.2%; Score 1663.4; DB 2; Length 2067; | | |
| Best Local Similarity | 99.9%; Pred. No. 0; | | |
| Matches 1664; Conservative | 0; Mismatches 1; Indels 0; Gaps 0; | | |
| Oy | 11 | ATGAGAGGGGACACCCACGGGAATGCAATCCAGCAAGAACACCTTCAGCTGAGCATCT | 70 |
| Db | 1 | ATGAGAGGGGACACCCACGGGAATGCAATCTCCAGCAAGAACACCTTCAGCTGAGCATCT | 60 |
| Oy | 71 | CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGGCCCAGGATCTCCAGCCCAAGCATCT | 130 |
| Db | 61 | CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGGCCCAGGATCTCCAGCCCAAGCATCT | 120 |
| Oy | 131 | CCAGCCCAAGGATCTCCAGCTGGGACACCTTCGGGGCCGGGATCTCCAGCCCAAGCATCT | 190 |

Db 121 CCAGCCAGGAGATCTCCAGCTGGGAGACCTCCGGGCGGGGATCTCCAGCCAGGATCT 180
 Qy 191 CCAGCTGGTACACCTCCAGCCCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 250
 Db 181 CCAGCTGGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 240
 Qy 251 CCAGCCCGGGGATCTCCGGCTCTGGGATCACTTTCCAGGCTCTCACTCCGAGGATCA 310
 Db 241 CCAGCCCGGGGATCTCCGGCTCTGGGATCACTTTCCAGGCTCTCACTCCGAGGATCA 300
 Qy 311 TCCGCGAGGATGAGCTCGGTGACAACTCCCAACAGAGTACCTTTGTAAGAACAA 370
 Db 301 TCCGCGAGGATGAGCTCGGTGACAACTCCCAACAGAGTACCTTTGTAAGAACAA 360
 Qy 371 CCAGTGGGGGGGTGATCCATCCGATCATCTCTGCGAGGTAGCAGCAGAACAGGACC 430
 Db 361 CCAGTGGGGGGGTGATCCATCCGATCATCTCTGCGAGGTAGCAGCAGAACAGGACC 420
 Qy 431 ACCAGGGAGAGCCCGAGGTACGAGGCTGCGCAAGTTACCTGGCGGGAGGCGCAAGAG 490
 Db 421 ACCAGGGAGAGCCCGAGGTACGAGGCTGCGCAAGTTACCTGGCGGGAGGCGCAAGAG 480
 Qy 491 CTACCGCTCATCGGGTGGGTCTCTCTCATTTGGCTTGGTGGTGGTGGTGGTGGTGGT 550
 Db 481 CTACCGCTCATCGGGTGGGTCTCTCTCATTTGGCTTGGTGGTGGTGGTGGTGGTGGT 540
 Qy 551 TTCCAGTTCTGGCAGGGGCGACACAGAGGTACAGGTACAGAGAGAGAGAGAGAGAGAG 610
 Db 541 TTCCAGTTCTGGCAGGGGCGACACAGAGGTACAGGTACAGAGAGAGAGAGAGAGAGAG 600
 Qy 611 AAGCAGCTGTTGGTGTGACGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 670
 Db 601 AAGCAGCTGTTGGTGTGACGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
 Qy 671 TGGCGAGGTTTGAAGTGGGACAAAGCTCTGCTTAAATCTACTCTGGGCTCTCCCATAG 730
 Db 661 TGGCGAGGTTTGAAGTGGGACAAAGCTCTGCTTAAATCTACTCTGGGCTCTCCCATAG 720
 Qy 731 TGGCTTCCCATCTGTAGAGAGCACTGGAATGACTCTTACTGAGAGAGAGAGAGAGAG 790
 Db 721 TGGCTTCCCATCTGTAGAGAGCACTGGAATGACTCTTACTGAGAGAGAGAGAGAGAG 780
 Qy 791 CTGGGTTTGAAGTGTCTACCGGACAAAGGTTGCCACAGGATTTTGGCAACAGC 850
 Db 781 CTGGGTTTGAAGTGTCTACCGGACAAAGGTTGCCACAGGATTTTGGCAACAGC 840
 Qy 851 TTCTCAATCTTGAAGTATCACTCCAGCTCCAGAGAAAGCTCCAGAGTCTGAATGCCCT 910
 Db 841 TTCTCAATCTTGAAGTATCACTCCAGCTCCAGAGAAAGCTCCAGAGTCTGAATGCCCT 900
 Qy 911 TCCGAGCGGATATCTCCCTCCAGGTTCCCACTGCGGAGTGAAGGAGCATGACCGGCGG 970
 Db 901 TCCGAGCGGATATCTCCCTCCAGGTTCCCACTGCGGAGTGAAGGAGCATGACCGGCGG 960
 Qy 971 ATCGTGGAGGGGCGCTGCTCGGATAGCAAGTGGCTTTGGCAAGTGAAGTCTGCACTTC 1030
 Db 961 ATCGTGGAGGGGCGCTGCTCGGATAGCAAGTGGCTTTGGCAAGTGAAGTCTGCACTTC 1020
 Qy 1031 GGGACCAACCAATCTGTGAGAGGACGCTCATTTGACGCGGAGTGGTGGTGGTGGTGGTGG 1090
 Db 1021 GGGACCAACCAATCTGTGAGAGGACGCTCATTTGACGCGGAGTGGTGGTGGTGGTGGTGG 1080
 Qy 1091 CACTGCTTCTTGTGACCCCGGGAGAGAGGTCTCTGAGGAGTGAAGGAGTGAAGGAGGAGC 1150
 Db 1081 CACTGCTTCTTGTGACCCCGGGAGAGAGGTCTCTGAGGAGTGAAGGAGTGAAGGAGGAGC 1140
 Qy 1151 AGCAACTGTGACCAAGTGGCTGTGAGGAGGCTTCATTTGCGAGATCATCAACAGCAAT 1210
 Db 1141 AGCAACTGTGACCAAGTGGCTGTGAGGAGGCTTCATTTGCGAGATCATCAACAGCAAT 1200
 Qy 1211 TACACCGAGTGAAGGAGAGCATATGACATCGCCCTCATGCGGCTGTCCAAAGCCCTGAGC 1270
 Db 1201 TACACCGAGTGAAGGAGAGCATATGACATCGCCCTCATGCGGCTGTCCAAAGCCCTGAGC 1260

Qy 1271 CTGTCCGCTCATCAACCTGTGCTTGGCTCCCATGATGAGACAGACTTTAGCTCAAT 1330
 Db 1261 CTGTCCGCTCATCAACCTGTGCTTGGCTCCCATGATGAGACAGACTTTAGCTCAAT 1320
 Qy 1331 GAGACTGTGCTGATCAAGGCTTTGGCAGACCGAGAGACAGATGACAGACATCCCCC 1390
 Db 1321 GAGACTGTGCTGATCAAGGCTTTGGCAGACCGAGAGACAGATGACAGACATCCCCC 1380
 Qy 1391 TTCTCCGGGAGGAGTGAAGTCAATCTGATGACCTTCAAGAAATGATGATCTAGTGGTC 1450
 Db 1381 TTCTCCGGGAGGAGTGAAGTCAATCTGATGACCTTCAAGAAATGATGATCTAGTGGTC 1440
 Qy 1451 TATGACAGTTACCTTACCCCAAGATGATGTGCTGGGAGCTTTGGTGGGAGAGAGAC 1510
 Db 1441 TATGACAGTTACCTTACCCCAAGATGATGTGCTGGGAGCTTTGGTGGGAGAGAGAC 1500
 Qy 1511 TCTGCGCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACAACTGGTACTGTG 1570
 Db 1501 TCTGCGCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACAACTGGTACTGTG 1560
 Qy 1571 GCAGGTGACACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACTGGTGTATCACCC 1630
 Db 1621 GCAGGTGACACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACTGGTGTATCACCC 1620
 Qy 1631 AAAGTGACAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAC 1675
 Db 1621 AAAGTGACAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAC 1665

RESULT 13
 BC018715
 LOCUS
 DEFINITION
 Homo sapiens transmembrane protease, serine 13, mRNA (cDNA clone IMAGE:3505123).
 VERSION
 BC018715.1 GI:19387970
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauener, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, A.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ueding, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loughellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwen, P.J., McKernan, K.D., Malek, U.A., Gunaratne, P.H., Richards, S., Wooley, K.C., Hale, S., Garcia, A.M., Gay, L.D., Hu, Y., Gibbs, R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Maman, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smolins, D.E., Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 CONSRPM
 TITLE
 JOURNAL

NIH MGC Project
 Direct Submission
 Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guan, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska,
Duane Smalls, Jeff Stolt, Miranda Tsai, George Yang, Jacques
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES

source

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 40 Row: a Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14042983.
Location/Qualifiers

1. 3222
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3505123"
/issue_type="Placenta, choriocarcinoma"
/clone_id="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 87.2%; Score 1523.4; DB 5; Length 3222;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
183 AGGATCTCCAGTGTGACCTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGCCC 242
Db ATGCATCTCCAGCAAGAACCTTCAGTGTGAGCATCTCCAGGCGGGCATCTCCAGCTG 128
243 AGGATCTCCAGGCGGGCATCTCCAGTGTGAGCATCTCCAGGCGGGCATCTCCAGGCA 302
129 GGAACCTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGGCA 188
303 GGTATCATCCGCGAGGTCAGCTCCGAGCAACCTCCCAACGAGTGTGATCTGTTA 362
189 GGTATCATCCGCGAGGTCAGCTCCGAGCAACCTCCCAACGAGTGTGATCTGTTA 248
363 GAGCAACCAAGTGGGGCTGTACCATCGATCATCTCTGCGAGGTCAAGCAGCA 422
249 GAGCAACCAAGTGGGGCTGTACCATCGATCATCTCTGCGAGGTCAAGCAGCA 308
423 CCAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 482
309 CCAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 368
483 AGAAGAGAGTACCGCTCATCGGGTGGGTCCTCTCTCATTTGCGTGTGTTTGGCTCA 542
369 AGAAGAGAGTACCGCTCATCGGGTGGGTCCTCTCTCATTTGCGTGTGTTTGGCTCA 428
543 TCATCTCTTCAGTTCTGCGAGGCGCACAGAGGATGAGTACAGAGAGAGAGAGAGAGAG 602
429 TCATCTCTTCAGTTCTGCGAGGCGCACAGAGGATGAGTACAGAGAGAGAGAGAGAGAG 488
603 GCTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
489 GCTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
663 AGCTGGGCTGGGTGAGTTGAGTGGAGCAAGTCTCTGCTTAAATCTACTCTGGGTCTCT 722
```

```
Db 549 AGCTGGGCTGGGTGAGTTGAGTGGAGCAAGTCTCTGCTTAAATCTACTCTGGGTCTCT 608
Qy 723 CCATTCAGTGGCTTTCCATCTGTAGAGCAATGGAATGACTCTACTCCAGAGAGAGCT 782
Db 609 CCATTCAGTGGCTTTCCATCTGTAGAGCAATGGAATGACTCTACTCCAGAGAGAGCT 668
Qy 783 GCCAGAGCTGGGTGAGAGTGTCTACCGGAGCAACCGAGGTTGCCAGAGAGATTTTG 842
Db 669 GCCAGAGCTGGGTGAGAGTGTCTACCGGAGCAACCGAGGTTGCCAGAGAGATTTTG 728
Qy 843 CCAAGAGCTTCTCATCTTGTAGATGAACTCCACCATTCAGAGAAAGCTTCACAGTCTG 902
Db 729 CCAAGAGCTTCTCATCTTGTAGATGAACTCCACCATTCAGAGAAAGCTTCACAGTCTG 788
Qy 903 AATGCTTCTTCCAGAGGATATCTCTCTCCAGTGTTCCTCCAGTGTTCCTCCAGTGTTC 962
Db 789 AATGCTTCTTCCAGAGGATATCTCTCTCCAGTGTTCCTCCAGTGTTCCTCCAGTGTTC 848
Qy 963 CCGGCGGATCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1022
Db 849 CCGGCGGATCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908
Qy 1023 TGCACTTGGAGCACCACCATCTGTGAGGAGCAGCTCATTTAGCGCCAGTGGTCTCA 1082
Db 909 TGCACTTGGAGCACCACCATCTGTGAGGAGCAGCTCATTTAGCGCCAGTGGTCTCA 968
Qy 1083 CTGCGGCGGATCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142
Db 969 CTGCGGCGGATCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1028
Qy 1143 CCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1202
Db 1029 CCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1088
Qy 1203 ACAGCAATTAACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262
Db 1089 ACAGCAATTAACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1148
Qy 1263 CCTGAGACCTGTCCGCTCAATCCACCTGTCTTCCCTCCAGAGAGAGAGAGAGAGAG 1322
Db 1149 CCTGAGACCTGTCCGCTCAATCCACCTGTCTTCCCTCCAGAGAGAGAGAGAGAGAG 1208
Qy 1323 GCGTCAATGAGAGCTGTGGATCAAGAGCTTTGGCAAGAGAGAGAGAGAGAGAGAGAG 1382
Db 1209 GCGTCAATGAGAGCTGTGGATCAAGAGCTTTGGCAAGAGAGAGAGAGAGAGAGAGAG 1268
Qy 1383 CATCCCTCTTCTCCGAGAGGAGTGAATCATCATGCACTTCAAGAAATGCAATGACT 1442
Db 1269 CATCCCTCTTCTCCGAGAGGAGTGAATCATCATGCACTTCAAGAAATGCAATGACT 1328
Qy 1443 ACTTGTCTATGAGATTAACCTTAACCAAGATGATGTGTGTGGGAGCTTGTGTGGG 1502
Db 1329 ACTTGTCTATGAGATTAACCTTAACCAAGATGATGTGTGTGGGAGCTTGTGTGGG 1388
Qy 1503 GCAGAGATCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1562
Db 1389 GCAGAGATCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1448
Qy 1563 GGTACTGTGAGAGGTGACCAAGCTGGGAGCAAGGCTGTGGCAGAGAAACCAACTGTG 1622
Db 1449 GGTACTGTGAGAGGTGACCAAGCTGGGAGCAAGGCTGTGGCAGAGAAACCAACTGTG 1508
Qy 1623 TGTACACCAAGTGAAGAGTTTCTTCCCTGTGATTTACAGCAAGATGAGAGAGAGTGC 1682
Db 1509 TGTACACCAAGTGAAGAGTTTCTTCCCTGTGATTTACAGCAAGATGAGAGAGAGTGC 1568
Qy 1683 GATTCAAGAAATCTTAACAGCTGGCTGTGCTGTGCTGACAGAGCAGGCTGTGTGA 1739
Db 1569 GATTCAAGAAATCTTAACAGCTGGCTGTGCTGTGCTGACAGAGCAGGCTGTGTGA 1625
```

RESULT 14
AB048797

LOCUS AB048797 3324 bp mRNA linear PRI 22-MAR-2001
DEFINITION Homo sapiens mps mRNA for mosaic serine protease, complete cds.
ACCESSION AB048797
VERSION AB048797.1 GI:13429971
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Kim,D.R., Sharmir,S., Inoue,M. and Kido,H.
TITLE Cloning and expression of novel mosaic serine proteases with and
JOURNAL without a transmembrane domain from human lung
PUBMED Biochim. Biophys. Acta 1518 (1-2), 204-209 (2001)
11267681
2 (bases 1 to 3324)
Kim,D.R., Inoue,M. and Kido,H.
AUTHORS Direct Submission
TITLE Submitted (12-SEP-2000) Hiroshi Kido, The University of Tokushima,
JOURNAL Division of Enzyme Chemistry, Institute for Enzyme Research,
Kuramoto-cho, Tokushima 770, Japan
(E-mail:kido@ier.tokushima-u.ac.jp, Tel.81-88-633-7424,
Fax:81-88-633-7425)
FEATURES
source Location/Qualifiers
1..3324
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="Lung"
/clone_lib="Marathon Ready cDNA (Clontech)"
1..3324
/gene="mps"
1..87
/gene="mps"
88..1701
/gene="mps"
/note="non-transmembrane type"
/codon_start=1
/transl_except=(pos:787..789,aa:Arg)
/transl_except=(pos:904..906,aa:His)
/transl_except=(pos:1498..1500,aa:His)
/product="mosaic serine protease"
/protein_id="BAB39742.1"
/db_xref="GI:13429972"
/translation="MERSHGNSAPRTSPSAGSPAQSPAGTPGRSPASPAQ
SPAGTPGRSPASPAQSPAGTPGRSPASPAQSPASPAQSPASPAQSPASPAQ
RASSGRSBSBASAVTTSPTRVIVRATPVCAVIRSPASATRAITRESPTVMO
GHTGIRKREORSCPKAVRCGVVDCLEKSDLDGCVRFMDKSLKITYSGSSHQMLP
ICSSNMNDSYSEKTRQROGFESAHRTTEVAHRDFANSFSLRVNSTIOESLHRSCHS
QRYISLOCSHGLRAMTGRIVGALASDKPMQVSLHPGTHICGGLIDIAQVLTAA
AHCFVTRKVLKGMKQVAGTSNHLPEASIAETIIINSYTDIEDVDIALRLK
PLTSAIHIPACIPMGQTFSLNETCWTGKTRRDTKSPILREVNQVLDLDFKRC
NDIYVYSYLTFRPMCAQDLHGKSDSCGDSGGPLVCBQNNRMYLAGTISNGTCGGR
NKPEVYKTVLEWLVLYSKMSVFRKRS"
1702..3324
/gene="mps"
3273..3278
/gene="mps"
ORIGIN
3' UTR
polyA_signal
Query Match 84.6%; Score 1479; DB 5; Length 3324;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1634; Conservative 0; Mismatches 0; Indels 135; Gaps 2;

QY 121 CCAGGATCTCCAGCCAGGAGATCTCCAGTGGAGACCTCCGGCCCGGATCTCCAGC 180
DB 198 CCAGGATCTCCAGCCAGGAGATCTCCAGTGGAGACCTCCGGCCCGGATCTCCAGC 257
QY 181 CCAGGATCTCCAGCTGGTACCTCCAGGCCGGGATCTCCAGGCCG----- 228
DB 258 CCAGGATCTCCAGCTGGTACCTCCAGGCCGGGATCTCCAGGCCGGGATCTCCAGC 317
QY 229 -----GGATCTCCAGCCAGGATCTCCAGGCCGGGATCTCCAGC 270
DB 318 CCAGGATCTCCAGCCAGGAGATCTCCAGGCCGGGATCTCCAGGCCGGGATCTCCAGC 377
QY 271 TCTGGATCACTTCCAGGTCCTTCATCCGAGGTATCATCCGACAGTCCGCTCGT 330
DB 378 TCTGGATCACTTCCAGGTCCTTCATCCGAGGTATCATCCGACAGTCCGCTCGT 437
QY 331 GACAACTCCCAACAGAGTGTACTTTTGAAGCAACCAATGGGGGTGTACCAT 390
DB 438 GACAACTCCCAACAGAGTGTACTTTTGAAGCAACCAATGGGGGTGTACCAT 497
QY 391 CCGATCATCTCTGTCAGGTCAGCAGCAGCAACAGAGGACCAAGAGAGCCAGGTAC 450
DB 498 CCGATCATCTCTGTCAGGTCAGCAGCAGCAACAGAGGACCAAGAGAGCCAGG--- 553
QY 451 GAGCTGCCCAAGTTCACTGGCGGAGGCGCAGAACAGCTACCGCTCATCGGTCGT 510
DB 554 ----- 553
QY 511 GCTCTCTCTCATTCGCTGGTGGTTTGGCTCATCTCTTCCAGTTTCCGAGGCCA 570
DB 554 -----TCAGTTTCCGAGGCCA 572
QY 571 CACAGGATCAGGTACAGAGAGCAGAGAGAGGTGTCCCAAGCACGCTGTCAGTGA 630
DB 573 CACAGGATCAGGTACAGAGAGCAGAGAGAGGTGTCCCAAGCACGCTGTCAGTGA 632
QY 631 CGGGGTGTGACTGCAAGCTGAAAGTGAAGACTGAGCTGGCTGCTGAGTTTGA 690
DB 633 CGGGGTGTGACTGCAAGCTGAAAGTGAAGACTGAGCTGGCTGCTGAGTTTGA 692
QY 691 CAACTCTCTGTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTACAG 750
DB 693 CAACTCTCTGTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTACAG 752
QY 751 CAACTGGAATGACTCTCTACTCAGAGAAAGCTGACAGCTGGTTTCAGAGATGCTCA 810
DB 753 CAACTGGAATGACTCTCTACTCAGAGAAAGCTGACAGCTGGTTTCAGAGATGCTCA 812
QY 811 CCGGACAAACGAGGTTGCCCAAGGATTTTCCCAAGCTTCTCAATCTTGAGATCAA 870
DB 813 CCGGACAAACGAGGTTGCCCAAGGATTTTCCCAAGCTTCTCAATCTTGAGATCAA 872
QY 871 CTCACCATCCAGAAAGCTTCCACAGTCTGAATGCTTCCAGAGGATATCTCCCT 930
DB 873 CTCACCATCCAGAAAGCTTCCACAGTCTGAATGCTTCCAGAGGATATCTCCCT 932
QY 931 CCAAGTTTCCCACTGCGGAGCTGAGAGGAGCATGACCGAGGATCGTGGAGAGGCGCTGCG 990
DB 933 CCAAGTTTCCCACTGCGGAGCTGAGAGGAGCATGACCGAGGATCGTGGAGAGGCGCTGCG 992
QY 1051 AGGACGCTCATTTGAGCGCCAGTGGGTGCTCATGCGCGCCCATGCTTTTCGAGACCG 1110
DB 1053 AGGACGCTCATTTGAGCGCCAGTGGGTGCTCATGCGCGCCCATGCTTTTCGAGACCG 1112
QY 1111 GGAAGAAGCTCTGGAAGGCTGGAAGGTGTACCGGAGCAGCAACCTGACCAAGTTGCC 1170
DB 1113 GGAAGAAGCTCTGGAAGGCTGGAAGGTGTACCGGAGCAGCAACCTGACCAAGTTGCC 1172
QY 1171 TGAGGAGCCTCATTCGCGAGATCATCAACAGCAATTACCGATGAGAGAGCA 1230

Db 1173 TGAGGAGCCTCCATTCGAGATCATCAAGCAATTACACGATAGAGAGCA 1232
QY 1231 CTATGATGAGCCTCATGAGGCTGTCCAAAGCCCTGACCTGTGCGTCAATCAACC 1290
Db 1233 CTATGATGAGCCTCATGAGGCTGTCCAAAGCCCTGACCTGTGCGTCAATCAACC 1292
QY 1291 TGCTGCTCCCAATGAGACGACCTTATGAGTGAAGACCTGAGTCAAGG 1350
Db 1293 TGCTGCTCCCAATGAGACGACCTTATGAGTGAAGACCTGAGTCAAGG 1352
QY 1351 CTTTGGCAAGACGAGGAGACAGATGACAAGACATCCCTCTCCGAGGAGTGACAGT 1410
Db 1353 CTTTGGCAAGACGAGGAGACAGATGACAAGACATCCCTCTCCGAGGAGTGACAGT 1412
QY 1411 CAATTCATGACTTCAAGAAATGCAATGACTACTTGTGTATGACAGTTACCTTACCCC 1470
Db 1413 CAATTCATGACTTCAAGAAATGCAATGACTACTTGTGTATGACAGTTACCTTACCCC 1472
QY 1471 AAGATGATGTGTGCTGAGGAGCCTTGTGAGGAGAGACCTCTCCAGGAGACAGCG 1530
Db 1473 AAGATGATGTGTGCTGAGGAGCCTTGTGAGGAGAGACCTCTCCAGGAGACAGCG 1532
QY 1531 GGGGCTCTTGTCTGTGAGCAAGAACCGCTGTACTGTGAGGAGTCAACAGCTGAGG 1590
Db 1533 GGGGCTCTTGTCTGTGAGCAAGAACCGCTGTACTGTGAGGAGTCAACAGCTGAGG 1592
QY 1591 CACAGGCTGTGAGGAGCAAGAACCGCTGTGTGTACACCAAGTGAACAGAGTTCTTCC 1650
Db 1593 CACAGGCTGTGAGGAGCAAGAACCGCTGTGTGTACACCAAGTGAACAGAGTTCTTCC 1652
QY 1651 CTGATTTTACAGCAAGATGAGAGCGAGTGCATTCAGAAATCTTAACTGAGTGGCT 1710
Db 1653 CTGATTTTACAGCAAGATGAGAGCGAGTGCATTCAGAAATCTTAACTGAGTGGCT 1712
QY 1711 GCTGCTCTGACAGACCGGCTGCTGTA 1739
Db 1713 GCTGCTCTGACAGACCGGCTGCTGTA 1741

RESULT 15
AX360089 1614 bp DNA linear PAT 13-FEB-2002
LOCUS AX360089
DEFINITION Sequence 45 from Patent W00200860.
ACCESSION AX360089
VERSION AX360089.1 GI:18675715
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS 1. Plowman, G., Whyte, D., Sudarsanam, S., Manning, G., Caenepeel, S. and
Charyczak, G.
TITLE Novel proteases
JOURNAL Patent: WO 0200860-A 45 03-JAN-2002;
Sugen, Inc. (US)
FEATURES
source 1. 1614
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 81.8%; Score 1429; DB 2; Length 1614;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 0; Indels 135; Gaps 2;
QY 11 ATGAGAGGAGGACAGACCGGAGATGATCTTCAGCAAGAACACCTTCAGTGAAGCATCT 70
Db 1 ATGAGAGGAGGACAGACCGGAGATGATCTTCAGCAAGAACACCTTCAGTGAAGCATCT 60

QY 71 CCAGCCAGGACATCTCAGCTGGAGACATCTCAGGCGGGCATCTCCAGCCAGGACATCT 130
Db 61 CCAGCCAGGACATCTCAGCTGGAGACATCTCAGGCGGGCATCTCCAGCCAGGACATCT 120
QY 131 CCAGCCAGGACATCTCAGCTGGAGACATCTCAGGCGGGCATCTCCAGCCAGGACATCT 190
Db 121 CCAGCCAGGACATCTCAGCTGGAGACATCTCAGGCGGGCATCTCCAGCCAGGACATCT 180
QY 191 CCAGCTGCTACCTCCAGGCGGGCATCTCCAGGCGG----- 228
Db 181 CCAGCTGCTACCTCCAGGCGGGCATCTCCAGGCGGCGGATCTCCAGCCAGGACATCT 240
QY 229 -----GGCATCTCAGCCAGGACATCTCAGGCGGGCATCTCCAGCTGAGATCA 280
Db 241 CCAGCCAGGACATCTCAGCTGGAGACATCTCAGGCGGGCATCTCCAGCTGAGATCA 300
QY 281 CTTTCCAGGCTCTCATCCGAGGATCATCATCCGAGTCAAGCTTCTGTTGACAACTCC 340
Db 301 CTTTCCAGGCTCTCATCCGAGGATCATCATCCGAGTCAAGCTTCTGTTGACAACTCC 360
QY 341 CCAGCCAGGATCTCTTGTGAGCAACCAAGTGGGCTGTACCATCCGATCATCT 400
Db 361 CCAGCCAGGATCTCTTGTGAGCAACCAAGTGGGCTGTACCATCCGATCATCT 420
QY 401 CTTGCGAGGTCAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 460
Db 421 CTTGCGAGGTCAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 466
QY 461 AAGTTCACTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 520
Db 467 ----- 466
QY 521 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
Db 467 -----TCAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 495
QY 581 AGGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 640
Db 496 AGGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 555
QY 641 GACTGCAAGCTGGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 700
Db 556 GACTGCAAGCTGGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 615
QY 701 CTTAAATCTAATCTGAGTCTCTCCATCAGTGTCTCTCCATCAGTGTCTCTCCATCAGT 675
Db 616 CTTAAATCTAATCTGAGTCTCTCCATCAGTGTCTCTCCATCAGTGTCTCTCCATCAGT 675
QY 761 GACTCTACTCAGAGAGACCTGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 820
Db 676 GACTCTACTCAGAGAGACCTGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 735
QY 821 GAGGTTGCCAGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 880
Db 736 GAGGTTGCCAGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 795
QY 881 CAGGAAGCTCCACAGAGTGTGAATGCTCTTCCAGGAGTGTATCTCTCTCTCAATGTTCC 940
Db 796 CAGGAAGCTCCACAGAGTGTGAATGCTCTTCCAGGAGTGTATCTCTCTCTCAATGTTCC 855
QY 941 CACTGCGGAGTGAAGGAGGAGTGAAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAG 1000
Db 856 CACTGCGGAGTGAAGGAGGAGTGAAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAG 915
QY 1001 AAGTGCCTTGGCAAGTGAAGTGTGACTTGGCAACCAACCAATCTGTGAGGAGGAGGAGT 1060
Db 916 AAGTGCCTTGGCAAGTGAAGTGTGACTTGGCAACCAACCAATCTGTGAGGAGGAGGAGT 975
QY 1061 ATTGAGCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1120
Db 976 ATTGAGCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1035
QY 1121 CTGAGAGGCTGGAAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1180

Db 1036 CTGGAGGAGCTGGAGGAGGTGTAGCGGGACCCAGCAACCTGCAACGATGGCTGAGGCGAGCC 1095
QY 1181 TCCATTGCCGAGATCATCATCAACAGCAATTAACCCGATGAGAGAGCACTATGACATC 1240
Db 1096 TCCATTGCCGAGATCATCATCAACAGCAATTAACCCGATGAGAGAGCACTATGACATC 1155
QY 1241 GCCCTCATGCGGCTGTCCAGCCCTGACCCCTGTCCGCTCAATCCATCCATCCCTGTGCTC 1300
Db 1156 GCCCTCATGCGGCTGTCCAGCCCTGACCCCTGTCCGCTCAATCCATCCATCCCTGTGCTC 1215
QY 1301 CCCATGATGAGACGACCTTTAGCTTCAATGAGACTGTGATCACAAGCTTTGGCAAG 1360
Db 1216 CCCATGATGAGACGACCTTTAGCTTCAATGAGACTGTGATCACAAGCTTTGGCAAG 1275
QY 1361 ACCAGGAGACAGATGACAAAGACATCCCTTCTCCGAGAGTGACAGGTCAATCTCATC 1420
Db 1276 ACCAGGAGACAGATGACAAAGACATCCCTTCTCCGAGAGTGACAGGTCAATCTCATC 1335
QY 1421 GACTTCAAGAAATGCAATGACTTACTGTCTATGACAGTTACTTACCCCAAGGATGATG 1480
Db 1336 GACTTCAAGAAATGCAATGACTTACTGTCTATGACAGTTACTTACCCCAAGGATGATG 1395
QY 1481 TGTGCTGGGGACCTTGTGGGGGAGAGACTCTGCGCAGGGAGACAGCGGGGGGCTCTT 1540
Db 1396 TGTGCTGGGGACCTTGTGGGGGAGAGACTCTGCGCAGGGAGACAGCGGGGGGCTCTT 1455
QY 1541 GTCTGTAGAGCAACAACCGCTGTACTGTGACAGGTGTCAACAGCTGGGGCACAGGCTGT 1600
Db 1456 GTCTGTAGAGCAACAACCGCTGTACTGTGACAGGTGTCAACAGCTGGGGCACAGGCTGT 1515
QY 1601 GGCCAGAGAAACAACCTGTGTGTACCAAGGTGACAGAGTTCTTCCCTGGAATTTAC 1660
Db 1516 GGCCAGAGAAACAACCTGTGTGTACCAAGGTGACAGAGTTCTTCCCTGGAATTTAC 1575
QY 1661 AGCAAGATGAGAGCGAGGTGCGATTGAGAAATCCTAA 1699
Db 1576 AGCAAGATGAGAGCGAGGTGCGATTGAGAAATCCTAA 1614

Search completed: September 18, 2006, 11:42:09
Job time : 9699 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 07:45:49 ; Search time 8720 Seconds
(without alignments)
11209.517 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 1748
Sequence: 1 ctccagacacatgagagaggg.....gctcgtcgtgacccgagaa 1748

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est3:*
- 3: gb_est4:*
- 4: gb_est5:*
- 5: gb_est6:*
- 6: gb_hic:*
- 7: gb_est2:*
- 8: gb_est7:*
- 9: gb_est8:*
- 10: gb_est9:*
- 11: gb_gss1:*
- 12: gb_gss2:*
- 13: gb_gss3:*
- 14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 1221.2 | 69.9 | 1472 | 6 | CR613669 full-leng |
| 2 | 1221.2 | 69.9 | 2029 | 4 | CR695643 full-leng |
| 3 | 969.4 | 55.5 | 1039 | 4 | BX356693 BX356693 |
| 4 | 813 | 46.5 | 821 | 4 | BX337410 BX337410 |
| 5 | 741 | 42.4 | 945 | 4 | BX337409 BX337409 |
| 6 | 711.4 | 40.7 | 930 | 3 | BU957076 AGENCOURT |
| 7 | 709.4 | 40.6 | 834 | 7 | BE732381 601569601 |
| 8 | 637.2 | 36.5 | 1043 | 3 | BU558176 AGENCOURT |
| 9 | 624 | 35.7 | 915 | 3 | BU860003 AGENCOURT |
| 10 | 621 | 35.5 | 693 | 3 | DN998208 TC106402 |
| 11 | 581.8 | 33.3 | 595 | 4 | CB270010 1008917 H |
| 12 | 552.8 | 31.6 | 777 | 4 | BX389632 BX389632 |
| 13 | 535.6 | 30.6 | 852 | 4 | DA830024 DA830024 |
| 14 | 532.8 | 30.5 | 640 | 4 | BX952562 DKFZP781X |
| 15 | 493.4 | 28.2 | 579 | 3 | BP359643 BP359643 |
| 16 | 486.2 | 27.8 | 845 | 9 | DN932137 AGENCOURT |
| 17 | 477.4 | 27.3 | 596 | 3 | BP319980 BP319980 |
| 18 | 471 | 26.9 | 878 | 4 | BX337187 BX337187 |
| 19 | 470.4 | 26.9 | 694 | 4 | CA113327 UT-CF-FNO |

| | | | | | | |
|----|-------|------|------|----|----------|----------|
| 20 | 468 | 26.8 | 583 | 3 | BP315554 | BP315554 |
| 21 | 468 | 26.8 | 590 | 9 | DB290424 | DB290424 |
| 22 | 462 | 26.4 | 586 | 9 | DA848528 | DA848528 |
| 23 | 453 | 25.9 | 575 | 9 | DB223460 | DB223460 |
| 24 | 452.8 | 25.9 | 579 | 9 | DA445009 | DA445009 |
| 25 | 449.8 | 25.7 | 583 | 3 | BP315772 | BP315772 |
| 26 | 440.8 | 25.2 | 678 | 9 | BE289529 | BE289529 |
| 27 | 427 | 24.4 | 748 | 9 | DN948363 | DN948363 |
| 28 | 411.6 | 23.5 | 614 | 7 | BE286322 | BE286322 |
| 29 | 386.4 | 22.1 | 679 | 5 | CP899959 | CP899959 |
| 30 | 359.2 | 20.5 | 377 | 5 | CD633581 | CD633581 |
| 31 | 358.8 | 20.5 | 622 | 7 | BF119003 | BF119003 |
| 32 | 357.6 | 20.5 | 378 | 5 | CD633580 | CD633580 |
| 33 | 355.4 | 20.3 | 364 | 8 | CN335419 | CN335419 |
| 34 | 348 | 19.9 | 402 | 10 | R78581 | R78581 |
| 35 | 335.8 | 19.2 | 621 | 7 | BE285038 | BE285038 |
| 36 | 298.8 | 17.1 | 1106 | 4 | BX356692 | BX356692 |
| 37 | 286 | 16.4 | 333 | 4 | CA406437 | CA406437 |
| 38 | 285 | 16.3 | 285 | 7 | AW845106 | AW845106 |
| 39 | 276.4 | 15.8 | 286 | 1 | A1909842 | A1909842 |
| 40 | 266.4 | 15.2 | 563 | 7 | BE290088 | BE290088 |
| 41 | 265.4 | 15.2 | 751 | 7 | BE280394 | BE280394 |
| 42 | 254.4 | 14.6 | 582 | 3 | BP356834 | BP356834 |
| 43 | 240.2 | 13.7 | 252 | 2 | BF757612 | BF757612 |
| 44 | 240.2 | 13.7 | 254 | 2 | BF757611 | BF757611 |
| 45 | 240.2 | 13.7 | 255 | 2 | BF757610 | BF757610 |

ALIGNMENTS

| | | | | | |
|-------------|--|---------|------|--------|-----------------|
| RESULT 1 | CR613669 | 1472 bp | mRNA | linear | HTC 21-JUL-2004 |
| LOCUS | full-length cDNA clone CSOD1041YM07 of Placenta Cot 25-normalized | | | | |
| DEFINITION | of Homo sapiens (human). | | | | |
| ACCESSION | CR613669 | | | | |
| VERSION | CR613669.1 GI:50494476 | | | | |
| KEYWORDS | HTC; CNSUT_cDNA. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | 1 (bases 1 to 1472) | | | | |
| AUTHORS | Li, W.B., Gruber, C., Jesse, J. and Polayes, D. | | | | |
| TITLE | Full-length cDNA libraries and normalization | | | | |
| JOURNAL | Unpublished | | | | |
| REMARK | Contact : Feng Liang Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue 2 (bases 1 to 1472) | | | | |
| REFERENCE | 2 (bases 1 to 1472) | | | | |
| AUTHORS | Genoscope. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqre@genoscope.cns.fr | | | | |
| COMMENT | - Web : www.genoscope.cns.fr - 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..1472 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /mol_type="mRNA" | | | | |
| | /db_xref="taxon:9606" | | | | |
| | /clone="CSOD1041YM07" | | | | |
| | /tissue type="Placenta Cot 25-normalized" | | | | |
| | /plasmid="pCMVSPORT_6" | | | | |
| ORIGIN | | | | | |
| Query Match | 69.9%; Score 1221.2; DB 6; Length 1472; | | | | |

Best Local Similarity 98.6%; Pred. No. 2e-281;
Matches 1248; Conservative 0; Mismatches 3; Indels 15; Gaps 1;

```

QY 31 GAATGCATCTCCAGAGAAACACCTTCAGCTGAGAGCATCTCCAGCCCGAGCATCTCCAGC 90
Db 1 GAATGCATCTCCAGAGAAACACCTTCAGCTGAGAGCATCTCCAGCCCGAGCATCTCCAGC 60
QY 91 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 150
Db 61 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 120
QY 151 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 210
Db 121 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 180
QY 211 CGGGGCACTCCAGAGCCG-----GGCATCTCCAGCCCGAGCATCTCCAGC 255
Db 181 CGGGGCACTCCAGAGCCGCGGAGCATCTCCAGCCCGAGCATCTCCAGC 240
QY 256 CGGGGCACTCTCCAGGCTCCAGCATCTTCAGAGGCTCTCATCCGAGAGTATCATCCGC 315
Db 241 CGGGGCACTCTCCAGGCTCCAGCATCTTCAGAGGCTCTCATCCGAGAGTATCATCCGC 300
QY 316 CAGGTCAAGCTTCGGTGAACCTCCCAACCAAGTGAATCTTTAGAGCAACCAAGT 375
Db 301 CAGGTCAAGCTTCGGTGAACCTCCCAACCAAGTGAATCTTTAGAGCAACCAAGT 360
QY 376 GGGGGCTGTACCATCCGATCATCTCCAGAGGTCAGACACAGCAACAGAGGCAACAG 435
Db 361 GGGGGCTGTACCATCCGATCATCTCCAGAGGTCAGACACAGCAACAGAGGCAACAG 420
QY 436 GAGAGCCAGAGTACGAGCTCCGCAAGTTACCTGGCGAGAGGCGCAAGAGCACTACC 495
Db 421 GAGAGCCAGAGTACGAGCTCCGCAAGTTACCTGGCGAGAGGCGCAAGAGCACTACC 480
QY 496 GCTCATCGAGGTCGCTGCTCTCTCTCATTTGCCCTGGTGGTTTCGCTCATCATCTCTTCA 555
Db 481 GCTCATCGAGGTCGCTGCTCTCTCTCATTTGCCCTGGTGGTTTCGCTCATCATCTCTTCA 540
QY 556 GTTCGAGAGGCGCAACAGAGGATCAGGTACAGAGGAGAGAGGAGAGCTGCCAACA 615
Db 541 GTTCGAGAGGCGCAACAGAGGATCAGGTACAGAGGAGAGAGGAGAGCTGCCAACA 600
QY 616 CGCTGTTGCTGTGAACGAGGTCGTGAAGCTGAAGAGTGAAGTGAAGTGAAGTGAAGT 675
Db 601 CGCTGTTGCTGTGAACGAGGTCGTGAAGCTGAAGAGTGAAGTGAAGTGAAGTGAAGT 660
QY 676 GAGGTTTGAAGTGAAGTCTCTGCTTAAATCTACTCTGAGTCTCTCCATCAAGTGGCT 735
Db 661 GAGGTTTGAAGTGAAGTCTCTGCTTAAATCTACTCTGAGTCTCTCCATCAAGTGGCT 720
QY 736 TCCCATCTGTAGAGCACTGGAATGATCTCCACTCAGAGAAAGCTGCCAGAGCTGG 795
Db 721 TCCCATCTGTAGAGCACTGGAATGATCTCTACTCAGAGAAAGCTGCCAGAGCTGG 780
QY 796 TTTGAGAGTGTCAACCGAGCAACCGAGGTTGCCCAAGAGATTTTGCACAAGCTTCTC 855
Db 781 TTTGAGAGTGTCAACCGAGCAACCGAGGTTGCCCAAGAGATTTTGCACAAGCTTCTC 840
QY 856 AATCTTGAGATACAACTCCACATTCAGAGAAAGCTCCACAGGTGTGAATGCCCTTCCA 915
Db 841 AATCTTGAGATACAACTCCACATTCAGAGAAAGCTCCACAGGTGTGAATGCCCTTCCA 900
QY 916 GCGGTATATCTCCCTCCAGAGTGTCCCACTGCGAGCTGAGAGGCAATGACCGAGGAGAT 975
Db 901 GCGGTATATCTCTCTCCAGAGTGTCCCACTGCGAGCTGAGAGGCAATGACCGAGGAGAT 960
QY 976 GAGAGAGGAGCTGAGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTGAAGTGAAGTGAAG 1035
Db 961 GAGAGAGGAGCTGAGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTGAAGTGAAGTGAAG 1020
QY 1036 CACCAACATCTGTGAGAGGAGCTGATGAGGCCAAGTGGTGTCTCATCTGCCGCCATGT 1095

```

```

Db 1021 CACCAACATCTGTGAGAGGAGCTGATGAGCGCCAGTGGGTGCTCATCTGCCGCCACTG 1080
QY 1096 CTTCTGCTGAGCCCGGAGAGAGTCTCTGAGAGGCTGGAAGGTGTACCGGAGCAACAGCA 1155
Db 1081 CTTCTGCTGAGCCCGGAGAGAGTCTCTGAGAGGCTGGAAGGTGTACCGGAGCAACAGCA 1140
QY 1156 CTTGACACAGTGGCTGAGAGGAGCTTCATTTCCGAGATCATCATCAACAGCAATTACAC 1215
Db 1141 CTTGACACAGTGGCTGAGAGGAGCTTCATTTCCGAGATCATCATCAACAGCAATTACAC 1200
QY 1216 CGATGAGAGAGAGCACTATGATCATTCGCGCTCATTCGCGCTGTCCAGAGCCCGAGACCTGTC 1275
Db 1201 CGATGAGAGAGAGCACTATGATCATTCGCGCTCATTCGCGCTGTCCAGAGCCCGAGACCTGTC 1260
QY 1276 CGCTCA 1281
Db 1261 CGGTGA 1266

```

RESULT 2
CR596543
LOCUS
DEFINITION full-length cDNA clone CS0D1015YU24 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR596543
VERSION CR596543.1 GI:50477350
KEYWORDS HTC; CNUST; CDNA.
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2029)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2029)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source 1..2029
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YU24"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 69.9%; Score 1221.2; DB 6; Length 2029;
Best Local Similarity 98.6%; Pred. No. 2.2e-281;
Matches 1248; Conservative 0; Mismatches 3; Indels 15; Gaps 1;

```

QY 31 GAATGCATCTCCAGAGAAACACCTTCAGCTGAGAGCATCTCCAGCCCGAGCATCTCCAGC 90
Db 1 GAATGCATCTCCAGAGAAACACCTTCAGCTGAGAGCATCTCCAGCCCGAGCATCTCCAGC 60
QY 91 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 150
Db 61 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 120
QY 151 TGGAGACCTTCAGAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGCCCGAGCATCTCCAGC 210

```

```

Db      121 TGGAGACCTCCGGGCGGGGATCTCCAGCCAGGACATCTCAGCTGGTACACTCCAGG 180
Qy      211 CCGGGGCACTTCACAGCCG-----GGCATCTCCAGCCAGGATCTCCAGC 255
Db      181 CCGGGGCACTTCACAGCCGCGGGGATCTCCAGCCAGGACATCTCCAGC 240
Qy      256 CCGGGGCACTTCAGGCTCGGATCACTTCCAGGCTCCATCCGAGAGGATCATCCGC 315
Db      241 CCGGGGCACTTCAGGCTCGGATCACTTCCAGGCTCCATCCGAGAGGATCATCCGC 300
Qy      316 CAGGTCAAGCTCGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 375
Db      301 CAGGTCAAGCTCGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 360
Qy      376 GGGGGCTGTACCTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 435
Db      361 GGGGGCTGTACCTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 420
Qy      436 GGAAGGCCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 495
Db      421 GGAAGGCCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 480
Qy      496 GGTCTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 555
Db      481 GGTCTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 540
Qy      556 GTTCTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 615
Db      541 GTTCTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 600
Qy      616 GGTCTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 675
Db      601 GGTCTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 660
Qy      676 GAGGTTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 735
Db      661 GAGGTTGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 720
Qy      736 TCCCATCTGTAGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 795
Db      721 TCCCATCTGTAGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 780
Qy      796 TTTTGAAGAGTCTCAAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 855
Db      781 TTTTGAAGAGTCTCAAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 840
Qy      856 AATCTTGAGATCACTCAAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 915
Db      841 AATCTTGAGATCACTCAAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 900
Qy      916 GCGGTATATCTCCCTCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 975
Db      901 GCGGTATATCTCCCTCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 960
Qy      976 GGGAGGGGGGCTGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1035
Db      961 GGGAGGGGGGCTGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1020
Qy      1036 CACCAACATCTGTAGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1095
Db      1021 CACCAACATCTGTAGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1080
Qy      1096 CTTTCTGTGAGCCGGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1155
Db      1081 CTTTCTGTGAGCCGGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1140
Qy      1156 CTTTCTGTGAGCCGGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1215
Db      1141 CTTTCTGTGAGCCGGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1200
Qy      1216 CGATGAGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1275

```

```

Db      1201 CGATGAGAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 1260
Qy      1276 CGGTCA 1281
Db      1261 CGGTCA 1266

RESULT 3
BX356693
LOCUS 1039 bp mRNA linear EST 08-APR-2004
DEFINITION BX356693 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSOD1015YL24 5-PRIME, mRNA sequence.
ACCESSION BX356693
VERSION BX356693.2 GI:46290274
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 1039)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30372079.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6066.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CSOD1015DF12QP1&c=6066.r.
FEATURES
Location/Qualifiers
source
1..1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1015YL24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 55.5%; Score 969.4; DB 4; Length 1039;
Best local Similarity 97.9%; Pred. No. 4.3e-221;
Matches 1002; Conservative 10; Mismatches 7; Indels 4; Gaps 3;
Qy      31 GAATGATCTCCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 90
Db      1 GAATGATCTCCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 60
Qy      91 TGGAGACCTTCAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 150
Db      61 TGGAGACCTTCAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 120
Qy      151 TGGAGACCTTCAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 210
Db      121 TGGAGACCTTCAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 180
Qy      211 CCGGGGCACTTCAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 270
Db      181 CCGGGGCACTTCAGGAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 240
Qy      271 TCTGGATCACTTCAGGATCACTCCAGGATCACTCCAGGATCACTCCAGGATCACTTGGAGCAACAGT 330

```

Db 241 TCTGGCATCACTTTCAGGTCCTCATCCGAGGTCATCATCCGCAAGTCAGCCTCGGT 300
 Oy 331 GACAACCTCCCAACAGAGTGTACTTGTAGAGCAACACAGTGGGGGCTGTACCAT 390
 Db 301 GACAACCTCCCAACAGAGTGTACTTGTAGAGCAACACAGTGGGGGCTGTACCAT 360
 Oy 391 CCGATCATCTCTGCAAGTCAGCAGCAACCAAGGCGCAACAGGAGAGCCAGGATAC 450
 Db 361 CCGATCATCTCTGCAAGTCAGCAGCAACCAAGGCGCAACAGGAGAGCCAGGATAC 420
 Oy 451 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGAGTACCGCTCATCTGGGTGCT 510
 Db 421 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGAGTACCGCTCATCTGGGTGCT 480
 Oy 511 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGAGAGGCGCA 570
 Db 481 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGAGAGGCGCA 540
 Oy 571 CACAGGATCAGGTACAAAGAGCAGAGGAGAGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 630
 Db 541 CACAGGATCAGGTACAAAGAGCAGAGGAGAGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 600
 Oy 631 CCGGGTGTGACCTGCAAGCTGAAAGAGTGAAGAGCTGGGCTGCGTGAAGTTTGAAGTGA 690
 Db 601 CCGGGTGTGACCTGCAAGCTGAAAGAGTGAAGAGCTGGGCTGCGTGAAGTTTGAAGTGA 660
 Oy 691 CAACTCTGCTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTAGAG 750
 Db 661 CAACTCTGCTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTAGAG 720
 Oy 751 CAACTGGAATGACTCTTACTAGAGAGAGAGCTGCCAGAGCTGGGTTTGAAGAGTCTCA 810
 Db 721 CAACTGGAATGACTCTTACTAGAGAGAGAGCTGCCAGAGCTGGGTTTGAAGAGTCTCA 780
 Oy 811 CCGAGCAACCGAGGTTGGCCACAGGAGTTTGGCCAAAGCTTCTCAATCTTGAAGTAA 870
 Db 781 CCGAGCAACCGAGGTTGGCCACAGGAGTTTGGCCAAAGCTTCTCAATCTTGAAGTAA 840
 Oy 871 CTCACACATCCAGGAAAGCTTCAAGAGTCTGAATGACCTTCCAGCGGATATCTCCCT 930
 Db 841 CTCACACATCCAGGAAAGCTTCAAGAGTCTGAATGACCTTCCAGCGGATATCTCCCT 900
 Oy 931 CCAAGTCTCCCACTGCGGACTGAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGC 990
 Db 901 CCAAGTCTCCCACTGCGGACTGAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGC 959
 Oy 991 CTCGGATGCAAGTGGCTTGGCAAGTGAAGTCTGACATTCGGCAACCAACCATCTGTGG 1050
 Db 960 CTCGGATGCAAGTGGCTTGGCAAGTGAAGTCTGACATTCGGCAACCAACCATCTGTGG 1016
 Oy 1051 AGG 1053
 Db 1017 GGG 1019

RESULT 4
 BX337410 821 bp mRNA linear EST 07-APR-2004
 LOCUS BX337410 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 DEFINITION clone CS0D1041Y07 5-PRIME, mRNA sequence.
 ACCESSION BX337410
 VERSION BX337410.2 GI:46272986
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1 (bases 1 to 821)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 2, 2003 this sequence version replaced gi:30339607.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with NotI and cloned
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 6066.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?cs=CS0D1041AG04QPlac=6066.r.
 Location/Qualifiers
 1..821
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1041Y07"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 46.5%; Score 813; DB 4; Length 821;
 Best Local Similarity 98.3%; Pred. No. 1,36-183;
 Matches 807; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

Db 31 GAATGATCTTCACAGAAACACCTTCAAGCTGAGCATCTTCAGGCCAGGATCTCCAGC 90
 Oy 1 GAATGATCTTCACAGAAACACCTTCAAGCTGAGCATCTTCAGGCCAGGATCTCCAGC 60
 Db 91 TGGGACACCTCCAGGCGGGGATCTCCAGGCCAGGATCTTCAGGCCAGGATCTCCAGC 150
 Oy 61 TGGGACACCTCCAGGCGGGGATCTCCAGGCCAGGATCTTCAGGCCAGGATCTCCAGC 120
 Db 151 TGGGACACCTCCAGGCGGGGATCTCCAGGCCAGGATCTTCAGGCCAGGATCTCCAGC 210
 Oy 121 TGGGACACCTCCAGGCGGGGATCTCCAGGCCAGGATCTTCAGGCCAGGATCTCCAGC 180
 Db 211 CCGGGCATCTTCAGAGCGGGGATCTCCAGGCCAGGATCTTCAGGCCAGGATCTCCAGC 270
 Oy 181 CCGGGCATCTTCAGAGCGGGGATCTCCAGGCCAGGATCTTCAGGCCAGGATCTCCAGC 240
 Db 271 TCTGGCATCACTTTCAGGTCCTCATCCGAGGTCATCATCGGCCAGGATCGAGCTGGT 330
 Oy 241 TCTGGCATCACTTTCAGGTCCTCATCCGAGGTCATCATCGGCCAGGATCGAGCTGGT 300
 Db 331 GACAACCTCCCAACAGAGTGTACTTGTAGAGCAACACAGTGGGGCTGTACCAT 390
 Oy 301 GACAACCTCCCAACAGAGTGTACTTGTAGAGCAACACAGTGGGGCTGTACCAT 360
 Db 391 CCGATCATCTCTGCAAGTCAGCAGCAACCAAGGCGCAACAGGAGAGCCAGGATAC 450
 Oy 361 CCGATCATCTCTGCAAGTCAGCAGCAACCAAGGCGCAACAGGAGAGCCAGGATAC 420
 Db 451 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGAGTACCGCTCATCTGGGTGCT 510
 Oy 421 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGAGTACCGCTCATCTGGGTGCT 480
 Db 511 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGAGAGGCGCA 570
 Oy 481 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGAGAGGCGCA 540
 Db 571 CACAGGATCAGGTACAAAGAGCAGAGGAGAGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 630
 Oy 541 CACAGGATCAGGTACAAAGAGCAGAGGAGAGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 600
 Db 631 CCGGGTGTGACCTGCAAGCTGAAAGAGTGAAGAGCTGGGCTGCGTGAAGTTTGAAGTGA 690
 Oy 601 CCGGGTGTGACCTGCAAGCTGAAAGAGTGAAGAGCTGGGCTGCGTGAAGTTTGAAGTGA 660

QY 691 CAAGTCTGCTTAAATCTACTGAGGCTCCCATGAGTGGCTTCCCATCTGAGCAG 750
 DB 661 CAAGTCTGCTTAAATCTACTGAGGCTCCCATGAGTGGCTTCCCATCTGAGCAG 720
 QY 751 CAAGTCTGCTTAAATCTACTGAGGCTCCCATGAGTGGCTTCCCATCTGAGCAG 810
 DB 721 CAAGTCTGCTTAAATCTACTGAGGCTCCCATGAGTGGCTTCCCATCTGAGCAG 780
 QY 811 CCGAGCAACCGAGGTTGCCACAGGATTTTGGCCACAGCT 851
 DB 781 CCGAGCAACCGAGGTTGCCACAGGATTTTGGCCACAGCT 821
 RESULT 5
 EX337409/c 945 bp mRNA linear EST 07-APR-2004
 LOCUS BX337409 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 DEFINITION clone CSOD1041YM07 3-PRIME, mRNA sequence.
 ACCESSION BX337409
 VERSION BX337409.2 GI:46271157
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 945)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization.
 Unpublished (2001)
 On May 2, 2003 this sequence version replaced gi:30337605.
 COMMENT
 Contact: Genoscope - Centre National de Sequencage
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 6066.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?b=CSOD1041AG04NP1&c=6066.r.
 FEATURES
 source
 1. 945
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1041YM07"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Query Match 42.4%; Score 741; DB 4; Length 945;
 Best Local Similarity 96.8%; Pred. No. 2,4e-166;
 Matches 758; Conservative 16; Mismatches 7; Indels 2; Gaps 2;

QY 679 GTTGAAGTGGACAAGTCTCTGCTTAAATCTACTGAGGCTCCCATGAGCAGT 738
 DB 765 GTTGAAGTGGACAAGTCTCTGCTTAAATCTACTGAGGCTCCCATGAGCAGT 706
 QY 739 CATCTGAGCAAGCACTGGAATGAATCTCTACTGAGAGAAAGCTCCAGCAGTGGTTT 798
 DB 705 CATCTGAGCAAGCACTGGAATGAATCTCTACTGAGAGAAAGCTCCAGCAGTGGTTT 646
 QY 799 CGAAGTCTCAACCGGACACCGAGGTTGCCACAGAGGATTTTGGCCACACTTCTCAAT 858
 DB 645 CGAAGTCTCAACCGGACACCGAGGTTGCCACAGAGGATTTTGGCCACACTTCTCAAT 586
 QY 859 CTTGAGATACAACTCCACATCCAGAAAGGCTCCACAGTCTGTAATCCCTCCACAGC 918
 DB 585 CTTGAGATACAACTCCACATCCAGAAAGGCTCCACAGTCTGTAATCCCTCCACAGC 526
 QY 919 GTATATCTCCCTCCAGTGTTCCTCCACTGAGAGCTGAGAGGCGGATGCTGGG 978
 DB 525 GTATATCTCCCTCCAGTGTTCCTCCACTGAGAGCTGAGAGGCGGATGCTGGG 466
 QY 979 AGGGCGCTGGCTTCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTCCGACAC 1038
 DB 465 AGGGCGCTGGCTTCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTCCGACAC 406
 QY 1039 CCACATCTGTGAGAGGACCGCTCATTTGACGCCAGTGGGTCTCACTGCGCCACTGCTT 1098
 DB 405 CCACATCTGTGAGAGGACCGCTCATTTGACGCCAGTGGGTCTCACTGCGCCACTGCTT 346
 QY 1099 CTTGCTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGTGTACGCGGACACAGCACT 1158
 DB 345 CTTGCTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGTGTACGCGGACACAGCACT 286
 QY 1159 GCACCAAGTTGCTTGAAGGACGCTTCATTTGCCGAGATCATCAACAGCAATTACACCGA 1218
 DB 285 GMAWAGTGTGATAGAGGACGCTCAATTTGCCGAGATCATCAACAGCAATTACACCGA 226
 QY 1219 TGAGAGGAGACATATGACATCGCCCTCATGCGGCTGTCGAAGCCCTGACCCCTGTCGG 1278
 DB 225 TGAGAGGAGACATATGACATCGCCCTCATGCGGCTGTCGAAGCCCTGACCCCTGTCGG 166
 QY 1279 TCA 1281
 DB 165 TGA 163
 RESULT 6
 BU957076 930 bp mRNA linear EST 21-OCT-2002
 LOCUS BU957076
 DEFINITION AGNCOURT 10615692 NIH_MGC_107 Homo sapiens cDNA clone
 IMAGE:6731027 5', mRNA sequence.
 ACCESSION BU957076
 BU957076.1 GI:24186636
 VERSION BU957076
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 930)
 NIH-MGC http://imgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: L1CM1056 row: m column: 10

FEATURES High quality sequence stop: 653.

Location/Qualifiers

1..930

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6731027"

/issue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC 107"

/note="Organ: breast; Vector: pOT87; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

40.7%; Score 711.4; DB 3; Length 930;

Best Local Similarity 97.6%; Pred. No. 3e-159; Matches 743; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

QY 236 CCAGCCCGGCGATCTCCAGCCCGGCGATCTCCGCTCTGGATCATCTTCCAGGCTCTCA 295

DB 1 CCAGCCCGGCGATCTCCAGCCCGGCGATCTCCGCTCTGGATCATCTTCCAGGCTCTCA 60

QY 296 TCCGCGAGTATCATCTCCGCGAGTCAAGCTCTGGTGAACAACCTCCCAACCAAGATGAC 355

DB 61 TCCGCGAGTATCATCTCCGCGAGTCAAGCTCTGGTGAACAACCTCCCAACCAAGATGAC 120

QY 356 CTTGTTAGAGCAACCAAGTGGGGCTGTACCATCCGATCATCTCTGCGAGGTCA 415

DB 121 CTTGTTAGAGCAACCAAGTGGGGCTGTACCATCCGATCATCTCTGCGAGGTCA 180

QY 416 CCAGCAACCAAGGCGCAACCAAGGAGAGCCCAAGTACAGAGCTTCCCAAGTCACTGCGG 475

DB 181 CCAGCAACCAAGGCGCAACCAAGGAGAGCCCAAGTACAGAGCTTCCCAAGTCACTGCGG 240

QY 476 GAGGCGCGAAGCAAGTACCGCTCATCGGGTGGCTCTCCCTCATTTGCCCTGTGTT 535

DB 241 GAGGCGCGAAGCAAGTACCGCTCATCGGGTGGCTCTCCCTCATTTGCCCTGTGTT 300

QY 536 TCGCTCATCATCTCTTCCAGTTCGAGGAGGCCACAGAGGATCAGGTACAAGAGCAG 595

DB 301 TCGCTCATCATCTCTTCCAGTTCGAGGAGGCCACAGAGGATCAGGTACAAGAGCAG 360

QY 596 AGGAGAGCTGTCCCAAGCAGCTGTGCTGTGAAGGGGTGTGTGACTGCAAGCTGAAG 655

DB 361 AGGAGAGCTGTCCCAAGCAGCTGTGCTGTGAAGGGGTGTGTGACTGCAAGCTGAAG 420

QY 656 AGTACGAGCTGGGTGGTGAAGTTTGAATGAGTCAAGTCTCTGCTTAAATCTACTCT 715

DB 421 AGTACGAGCTGGGTGGTGAAGTTTGAATGAGTCAAGTCTCTGCTTAAATCTACTCT 480

QY 716 GGGTCTCTCCATCAGTGGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAG 775

DB 481 GGGTCTCTCCATCAGTGGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAG 540

QY 776 AAGACTCTCCAGCAGCTGGGTTTGAAGTGTCTACCGGACCAACCGAGGTTGCCACAGG 835

DB 541 AAGACTCTCCAGCAGCTGGGTTTGAAGTGTCTACCGGACCAACCGAGGTTGCCACAGG 600

QY 836 GATTTTGGCAACAGTTTCTCAATCTTGAATACAACTCAACATCCAGAGAAAGCTCTCAC 895

DB 601 GATTTTGGCAACAGTTTCTCAATCTTGAATACAACTCAACATCCAGAGAAAGCTCTCAC 660

QY 896 AGGTCTGAATGCCCTTCCAGAGGATATCT-CCCTCCAGTGTTCCTCACT-GCGGACTGA 953

DB 661 AGGTCTGAATGCCCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCACTGCGGACTGA 720

QY 954 GGGCCATGACCGGGCGGATCTGTGGAGGGGCGCTGGCTCG 994

Db 721 GGGCCATGACCGGGCGGATCTGTGGAGGGGCGCTGG 761

RESULT 7 BE732381 834 bp mRNA linear EST 15-SEP-2000

LOCUS BE732381

DEFINITION 60156960.F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844129 5', mRNA sequence.

ACCESSION BE732381

VERSION BE732381.1 GI:10146373

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 834)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov

Plate: LUCM540 Row: D Column: 02

High quality sequence stop: 817.

FEATURES

Location/Qualifiers

1..834

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3844129"

/issue_type="choriocarcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC 21"

/note="Organ: placenta; Vector: pOT87; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

40.6%; Score 709.4; DB 7; Length 834;

Best Local Similarity 95.4%; Pred. No. 8.9e-159; Matches 795; Conservative 0; Mismatches 31; Indels 7; Gaps 6;

QY 664 GCTGGGCTGCGTGAAGTTTGAATGAGTCTCTGCTTAAATCTACTTGGGTCTC 723

DB 1 GCTGGGCTGCGTGAAGTTTGAATGAGTCTCTGCTTAAATCTACTTGGGTCTC 60

QY 724 CCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAATGACTCTCACTCAGAGAGACTG 783

DB 61 CCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAATGACTCTCACTCAGAGAGACTG 120

QY 784 CCAGCAGCTGGGTTTGAAGTGTCTCAACCGGACCAACGAGGTTGCCACAGGATTTTGC 843

DB 121 CCAGCAGCTGGGTTTGAAGTGTCTCAACCGGACCAACGAGGTTGCCACAGGATTTTGC 180

QY 844 CAACAGCTTCAATCTTGAATACAACTCCACATCCAGAGAAAGCTCTCACAGGTTGGA 903

DB 181 CAACAGCTTCAATCTTGAATACAACTCCACATCCAGAGAAAGCTCTCACAGGTTGGA 240

QY 904 ATGCCCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCCACTGCGGAGCTGAGAGGCGATGAC 963

DB 241 ATGCCCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCCACTGCGGAGCTGAGAGGCGATGAC 300

QY 964 CCGGCGATCGGAGGGGGGCTGGCTCGGATGAGCAAGTGGCTTGGCAAGTACT 1023
DB 301 CCGGCGATCGTGGAGGGGGGCTGGCTCGGATGAGCAAGTGGCTTGGCAAGTACT 360
QY 1024 GCACCTTGGGCAACCAACCATCTGTGAGGAGCAGCTCATTTGACGCCAGTGGTGTCTAC 1083
DB 361 GCACCTTGGGCAACCAACCATCTGTGAGGAGCAGCTCATTTGACGCCAGTGGTGTCTAC 420
QY 1084 TCCGCGCCACTGCTTCTTCTGAGACCCGGGAGAGAGTCTTGGAGGGCTGGAAGGTATGC 1143
DB 421 TCCGCGCCACTGCTTCTTCTGAGACCCGGGAGAGAGTCTTGGAGGGCTGGAAGGTATGC 480
QY 1144 GGGGACGAGCAACCTGACCAAGTTCCTGAGGAGAGCTTCATTTGGCCAGATCATCAAA 1203
DB 481 GGGGACGAGCAACCTGACCAAGTTCCTGAGGAGAGCTTCATTTGGCCAGATCATCAAA 538
QY 1204 CAGCAATTACACCGATGAGAGAGAGCACTA-TGACATGAGCTTCATTTGGGCTTCCAGC 1262
DB 539 CAGCAATTACACCGATGAGAGAGAGCACTAATTGACATGAGCTTCATTTGGGCTTCCAG 598
QY 1263 CCTGACCTCTGCTGCTGATCATTCACCTGCTGCTT-CCCCATGATGAGACAGCTTT 1321
DB 599 CCCCCTGAACTGTCTGCTGATCATTCACCTGCTGCTTCCCCATGATGAGACAGCTTT 658
QY 1322 AGCCTCATGAGACCTGCTGATCATCAGGCTTTGGCAAGAGGAGAGACAGATGACAG 1381
DB 659 AGCCTCATGAGACCTGCTGATCATCAGGCTTTGGCAAGAGGAGAGACAGATGACAG 718
QY 1382 ACATCCCCCTTCTCTCCGGAGGTGAGGTCAATCTCATGCACTTCAGAAATGCAATGAC 1441
DB 719 ACATCCCCCTTCTCTCCGGAGGTGAGGTCAATCTCATGCACTTCAGAAATGCAATGAC 777
QY 1442 TACTTGTCTATGACAGTTACCTTAACCCAGATGATGTGTGCTGGGAGCT 1494
DB 778 TAACTGTCTATGACAGT--ACCTTAACCCAGATGATGTGTGCTGGGAGCT 828

RESULT 8
LOCUS BUS58176 1043 bp mRNA linear EST 16-SEP-2002
DEFINITION AGENCOURT 10212529 NIH MGC 107 Homo sapiens cDNA clone
IMAGE:6586431 5', mRNA sequence.
BUS58176
ACCESSION BUS58176.1 GI:22908472
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1043)
AUTHORS NIH-MGC htp://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsab@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2799 row: P column: 15
High quality sequence stop: 489.
Location/Qualifiers
1..1043
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6586431"
/tissue_type="adenocarcinoma, cell line"

ORIGIN
Query Match 36.5%; Score 637.2; DB 3; Length 1043;
Best Local Similarity 98.6%; Pred. No. 1.9e-141;
Matches 653; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
/lab host="DH10B (phage-resistant)"
/clone lib="NIH MGC 107"
/note:Organ: breast; Vector: pORF7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT printing.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

QY 236 CCAGCCAGGATCTCAGAGCCGGGATCTCCGCTCTGGCATCATTTCCAGGTCTCA 295
DB 1 CCAGCCAGGATCTCAGAGCCGGGATCTCCGCTCTGGCATCATTTCCAGGTCTCA 60
QY 296 TCCGCGAGTGCATATCGGCGAGGTCAGGCTCGGTCAGAACCTCCCAACAGAGTAC 355
DB 61 TCCGCGAGTGCATATCGGCGAGGTCAGGCTCGGTCAGAACCTCCCAACAGAGTAC 120
QY 356 CTGTGAGCAACACAGATGGGGCTGTACCCATCCGATCATCTCTGCGAGGTACAGA 415
DB 121 CTGTGAGCAACACAGATGGGGCTGTACCCATCCGATCATCTCTGCGAGGTACAGA 180
QY 416 CCAGCAACAGAGGTCACAGGAGAGAGCCAGTACAGAGCTGCGCAAGTCACTGGCGG 475
DB 181 CCAGCAACAGAGGTCACAGGAGAGAGCCAGTACAGAGCTGCGCAAGTCACTGGCGG 240
QY 476 GAGGCGCAAGAGAGTACAGGCTCATGCGGTCGCTCCCTCATTTGCCCTGGTGT 535
DB 241 GAGGCGCAAGAGAGTACAGGCTCATGCGGTCGCTCCCTCATTTGCCCTGGTGT 300
QY 536 TCGCTCATCATCTCTTCCAGTCTGTGAGGAGCCACAGAGATCAGTACAGAGAG 595
DB 301 TCGCTCATCATCTCTTCCAGTCTGTGAGGAGCCACAGAGATCAGTACAGAGAG 360
QY 596 AGGAGAGAGTGTCCCAAGACAGCTGTTGCTGTGACGGGGTGTGACCTGCAAGTGA 655
DB 361 AGGAGAGAGTGTCCCAAGACAGCTGTTGCTGTGACGGGGTGTGACCTGCAAGTGA 420
QY 656 AGTACAGAGTGTGCTGAGAGTGTGACCTGAGGACAGTCTGCTTAATTAATTA 715
DB 421 AGTACAGAGTGTGCTGAGAGTGTGACCTGAGGACAGTCTGCTTAATTAATTA 480
QY 716 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGAGCAAGTGAATGATCTCTACAGAG 775
DB 481 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGAGCAAGTGAATGATCTCTACAGAG 540
QY 776 AAGACTGCGCAGAGCTGGGTTTGAAGTGTGACCGGAGACCGAGAGTGGCCACAG 835
DB 541 AAGACTGCGCAGAGCTGGGTTTGAAGTGTGACCGGAGACCGAGAGTGGCCACAG 600
QY 836 GATTTGCAACAGCTTCTCAA-TCTTGAGATAACTCCACCATCCAGGAAACCTCCA 894
DB 601 GATTTGCAACAGCTTCTCAA-TCTTGAGATAACTCCACCATCCAGGAAACCTCCA 660
QY 895 CA 896
DB 661 CA 662

RESULT 9
LOCUS BUS60003 915 bp mRNA linear EST 16-OCT-2002
DEFINITION AGENCOURT 10435147 NIH MGC 107 Homo sapiens cDNA clone
IMAGE:6651006 5', mRNA sequence.
BUS60003
ACCESSION BUS60003.1 GI:24044995
VERSION EST.
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo;
1 (bases 1 to 915)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsbgs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNCM2897 row: c column: 06
High quality sequence atop: 460.
Location/Qualifiers
1..915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:665106"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 107"
/note="Organ: breast; Vector: pOT87; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

ORIGIN
Query Match 35.7%; Score 624; DB 3; Length 915;
Best Local Similarity 98.2%; Pred. No. 2.7e-138;
Matches 641; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

236 CCAGCCGAGGATCTCCAGCCCGGAGCTCTCGGCTTGAGATCACTTCAGAGTCTCA 295
1 CCAGCCGAGGATCTCCAGCCCGGAGCTCTCGGCTTGAGATCACTTCAGAGTCTCA 60
236 TCCGCGAGGTATCTCCGCGAGGTCTGAGTGAACAACCTCCCAACGAGTGTAC 355
61 TCCGCGAGGTATCTCCGCGAGGTCTGAGTGAACAACCTCCCAACGAGTGTAC 120
356 CTTGTTAGAGCAACAGAGTGGGGCTGTACCATCCGATCATCTCTCCAGGTACAGA 415
121 CTTGTTAGAGCAACAGAGTGGGGCTGTACCATCCGATCATCTCTCCAGGTACAGA 180
416 CCAGCAACGAGGCGCAGGAGGAGGCCAGGTACAGCTTGCACCAAGTCACTTGGCGG 475
181 CCAGCAACGAGGCGCAGGAGGAGGCCAGGTACAGCTTGCACCAAGTCACTTGGCGG 240
476 GAGGCGCAGGAGGAGCTACCGCTCATCGGGTGGCTCTCCATTGACCTGTGTT 535
241 GAGGCGCAGGAGGAGCTACCGCTCATCGGGTGGCTCTCCATTGACCTGTGTT 300
536 TCGCTCATCATCTCTTCAGATTCTGAGAGGGCCACAGAGGATCAGGTACAAGAGAG 595
301 TCGCTCATCATCTCTTCAGATTCTGAGAGGGCCACAGAGGATCAGGTACAAGAGAG 360
596 AGGAGAGCTGTCCCAAGACAGCTGTTCCTGTGAAGGGGTGTGAGTCAAGCTGAAG 655
361 AGGAGAGCTGTCCCAAGACAGCTGTTCCTGTGAAGGGGTGTGAGTCAAGCTGAAG 420
656 AGTACGAGCTGGGTGGGTGAGTTTGAAGTGGAGCAAGTCTCTGCTTAAATCTACTCT 715

Db 421 AGTACGAGCTGGGCTGGTGAGGTGTGACTGGGACAAGTCTCTGCTTAAATCTACTCT 480
Oy 716 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAG 775
Db 481 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAG 540
Oy 776 AAGACTTCCAGCAGCTGGGTTTGAAGTGTCTACCGGACACCGAGTTGCCACAC-G 834
Db 541 AAGACTTCCAGCAGCTGGGTTTGAAGTGTCTACCGGACACCGAGTTGCCACACAG 600
Oy 835 GGATTTTCCCAAGCTTCTCATCTTGAGTACAGCTCCACCATCCAGGAAA 887
Db 601 GGATTTTCCCAAGCTTCTCATCTTGAGTACAGCTCCACCATCCAGGAAA 653

RESULT 10
DN998208
LOCUS
DEFINITION DN998208 693 bp mRNA linear EST 17-MAY-2005
TC106402 Human breast cancer tissue, large insert, pCMV expression
library Homo sapiens cDNA clone TC106402 5' similar to Homo sapiens
transmembrane protease, serine 13 (TMPRSS13), mRNA sequence.
ACCESSION DN998208
VERSION DN998208.1 GI:66258035
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo;
1 (bases 1 to 693)
REFERENCE Birkebeck, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L.,
Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M.,
Zhang, X., Jay, G., and He, W.
High-throughput cloning of full-length human cDNAs directly from
cDNA libraries optimized for large and rare transcripts
unpublished (2005)
CONTACT: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com
This EST submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 Spine forward vector primer, Origene
Technologies Inc.
FEATURES
SOURCE
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC106402"
/tissue_type="Breast Cancer"
/clone_id="Human breast cancer tissue, large insert, pCMV
expression library"
/note="Organ: Mammary gland (cancer tissue); Vector:
pCMV6-Xis; Site_1: EcoRI; Site_2: XhoI/Sall compatible and
ligatio: oligo-dT primed reverse transcription optimized
for large and GC rich mRNA transcripts, cDNA size
selection, optimized ligator for large inserts into
mammalian expression vector, random clones selected for
end sequence verification of full-length genes"

ORIGIN
Query Match 35.5%; Score 621; DB 9; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.3e-137;

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30462916.
Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6066.X
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BAF026ZE03_AF02422_1&c=6066.x

FEATURES

Source Location/Qualifiers
1..777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1041Y07"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 31.6%; Score 552.8; DB 4; Length 777;
Best Local Similarity 98.6%; Pred. No. 3.1e-121;
Matches 557; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

717 GATCTCCATCAGAGGCTTCCATCTGTAGAGCAGCATGGAATGACTCTTACTGAGGA 776
Db 4 GGGTCTCCATAGTGGCTTCCATCTGTAGAGCAGCATGGAATGACTCTTACTGAGGA 63
Qy 777 AGACCTGCAGCAGCTGGGTTTCGAGAGTGCTCACCGGACACCGAGGTTGCCACAGG 836
Db 64 AGACCTGCAGCAGCTGGGTTTCGAGAGTGCTCACCGGACACCGAGGTTGCCACAGG 123
Qy 837 ATTTTGCACAGCTTCTCAATCTTGAGATACAACTCCACCAATCCAGAAAGCTCCACA 896
Db 124 ATTTTGCACAGCTTCTCAATCTTGAGATACAACTCCACCAATCCAGAAAGCTCCACA 183
Qy 897 GATCGATGCGCTTCCAGAGGATATCTCCCTCCAGTGTCCACCTGCGGAGCTGAAGG 956
Db 184 GATCGATGCGCTTCCAGAGGATATCTCCCTCCAGTGTCCACCTGCGGAGCTGAAGG 243
Qy 957 CCATGACCGGCGGATCGTGGAGAGGGGCTGCTCCGATAGCAAGTGCGCTTGCAAG 1016
Db 244 CCATGACCGGCGGATCGTGGAGAGGGGCTGCTCCGATAGCAAGTGCGCTTGCAAG 303
Qy 1017 TGAAGTCTGCACTTGGGACCAACCATCTGTGAGGACCGCTCATTTAGCGCCAGTGG 1076
Db 304 TGAAGTCTGCACTTGGGACCAACCATCTGTGAGGACCGCTCATTTAGCGCCAGTGG 363
Qy 1077 TGCTCACTGCGGCGGATCTTCTTGTGACCCCGGAGAAAGTCTTGAAGGGCTGGAAG 1156
Db 364 TGCTCACTGCGGCGGATCTTCTTGTGACCCCGGAGAAAGTCTTGAAGGGCTGGAAG 423
Qy 1137 TGTACGCGGGGACACGACCACTGACAGTTGCTTGAGGAGCGCTTCATGCCGATCA 1196
Db 424 TGTACGCGGGGACACGACCACTGACAGTTGCTTGAGGAGCGCTTCATGCCGATCA 483
Qy 1197 TCATCAACAGCAATTAACCGATGAGAGAGAGTATGACATGCGCTCATGCGGCTGT 1256
Db 484 TCATCAACAGCAATTAACCGATGAGAGAGAGTATGACATGCGCTCATGCGGCTGT 543
Qy 1257 CCAAGCCCTGACCGCTGCGGCTCA 1281
Db 544 CCAAGCCCTGACCGCTGCGGCTCA 568

RESULT 13
DB30024
LOCUS
DEFINITION DB30024 PLAC1 Homo sapiens cDNA clone PLAC1004270 5', mRNA
sequence.
ACCESSION DB30024
VERSION DB30024.1 GI:82330405
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)

Homo sapiens (human)

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
1 (bases 1 to 852)

REFERENCE

Kimura, K., Makamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Isuriant, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Matsushima, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560

COMMENT

CONTACT: Takao Isogai
Fuj Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: fuj-cdna@hriky.com
HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
Research Institute (HRI); cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
HRI.

FEATURES

Source Location/Qualifiers
1..852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC104270"
/issue_type="Placenta"
/clone_lib="PLAC1"
/note="Vector: pME18FLJ3"

ORIGIN

Query Match 30.6%; Score 535.6; DB 9; Length 852;
Best Local Similarity 93.1%; Pred. No. 4.2e-117;
Matches 632; Conservative 0; Mismatches 22; Indels 25; Gaps 6;

1 CTCGAGACCTTGAGAGGACACGACCGGAAATGCAATCTCCACAGAAACCTTCAGC 60
Db 81 CTCGAGACCTTGAGAGGACACGACCGGAAATGCAATCTCCACAGAAACCTTCAGC 140
Qy 61 TGAAGCATCTCCAGCGGACCTTCAGCTGGGACACCTCCAGGCGGGCATCTCCAGC 120
Db 141 TGAAGCATCTCCAGCGGACCTTCAGCTGGGACACCTCCAGGCGGGCATCTCCAGC 200
Qy 121 CCAAGCATCTCCAGCGGACCTTCAGCTGGGACACCTCCAGGCGGGCATCTCCAGC 180
Db 201 CCAAGCATCTCCAGCGGACCTTCAGCTGGGACACCTCCAGGCGGGCATCTCCAGC 260
Qy 181 CCAAGCATCTCCAGCTGGTACACCTCCAGGCGGGGATCTTCAGGCGG----- 228
Db 261 CCAAGCATCTCCAGCTGGTACACCTCCAGGCGGGGATCTTCAGGCGGCGGATCTCCAGC 320
Qy 229 ---GGCATCTCCAGCGGACCTTCAGCGCGGGGATCTTCGGGCTTGAGCATCACTTC 285
Db 321 CCAAGCATCTCCAGCGGACCTTCAGCGCGGGGATCTTCGGGCTTGAGCATCACTTC 380

QY 286 CAGGTCCTCATCCGCGAGTCATCATCCGAGGTGACGCTCGGTGACAACTCCCCAC 345
 |||||
 Db 381 CAGGTCCTCATCCGCGAGTCATCATCCGAGGTGACGCTCGGTGACAACTCCCCAC 440
 |||||
 QY 346 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCTCCATCCGATCATCTCTGC 405
 |||||
 Db 441 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCTCCATCCGATCATCTCTGC 500
 |||||
 QY 406 CAGGTCAGACAC-AGCAACCAAGGGCCACAGAGAGAGCCCAAGTACG-AGCTGCCCAAG 463
 |||||
 Db 501 CAGGTCAGACACCAAGCAACCAAGGGCCACAGAGAGAGCCCAAGTACGAAAGCTGCCCAAG 560
 |||||
 QY 464 TTCACTGCGCGAGAGGCGC-AGAAAGAGCTACCGCTCATCGGAGCGCTCTCTCTCAT 522
 |||||
 Db 561 TTCACTGCGCGAGAGGCGCAGAAAGAGCTACCGCTCATCGGAGCGCTCTCTCTCAT 620
 |||||
 QY 523 TGCCCTGGTGTTCGCTCATCTCTCTTCAGTTCTGGAGGGCCACAGAGGATCAG 582
 |||||
 Db 621 TGCCCTGGTGTTCGCTCATCTCTCTTCAGTTCTGGAGGGCCACAGAGGATCAG 680
 |||||
 QY 583 GTACAGAGACAGAGGAGAGCTGT---CCCAAGACGCTGT---CGCTGTACCGGG 635
 |||||
 Db 681 GTACAGAGACCAAGAGGAGAGCTGTCCAAAGCACGCTGTTCGCTGTGAAACGCGGT 740
 |||||
 QY 636 TGTGTGACTGCAGAGCTGAA 654
 |||||
 Db 741 GGTGTGACTGCAGAGCTGGA 759
 |||||

RESULT 14
 BX952562 640 bp mRNA linear EST 01-MAR-2004
 LOCUS DKZP781K14199.F1.781 (synonym: hlc4) Homo sapiens cDNA clone
 DEFINITION DKZP781K14199.5, mRNA sequence.
 ACCESSION BX952562
 VERSION BX952562.1 GI:43429741
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Oeanger, A.,
 Fodor, G., Han, M., and Wiemann, S.
 TITLE EST (Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., et al.)
 JOURNAL Unpublished (2003)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFZp781K14199) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcentzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers
 1..640
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp781K14199"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_id="781 (synonym: hlc4)"
 /note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI; B;
 cDNA-collection"

ORIGIN
 Query Match 30.5%; Score 532.8; DB 4; Length 640;

Best Local Similarity 97.0%; Pred. No. 1,86-116;
 Matches 559; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 CTCGAGACCATGAGAGGACACAGCCACGGGAATGCACTTCCACCAAGAACCTTACG 60
 |||||
 Db 65 CTCGAGACCATGAGAGGACACAGCCACGGGAATGCACTTCCACCAAGAACCTTACG 124
 |||||
 QY 61 TGGAGCATCTTCACGCCAGGATCTCCAGCTGGAGACCTCCAGGCCGGGATCTCCAGC 120
 |||||
 Db 125 TGGAGCATCTTCACGCCAGGATCTCCAGCTGGAGACCTCCAGGCCGGGATCTCCAGC 184
 |||||
 QY 121 CCAGGCACTTCCAGCCCAAGGATCTCCAGCTGGAGACCTCCGGGCCGGGATCTCCAGC 180
 |||||
 Db 185 CCAGGCACTTCCAGCCCAAGGATCTCCAGCTGGAGACCTCCGGGCCGGGATCTCCAGC 244
 |||||
 QY 181 CCAGGCACTTCCAGCTGGATCACTCCAGGCCGGGATCTCCAGCCG-----228
 |||||
 Db 245 CCAGGCACTTCCAGCTGGATCACTCCAGGCCGGGATCTCCAGCCGCGGATCTCCAGC 304
 |||||
 QY 229 ---GGCATCTTCACGCCAGGATCTCCAGCCGGGATCTCCGGCTTGGATCACTTTC 285
 |||||
 Db 305 CCAGGCACTTTCACGCCAGGATCTCCAGCCGGGATCTCCGGCTTGGATCACTTTC 364
 |||||
 QY 286 CAGGTCCTCATCCGCGAGTCATCATCCGAGGTGACGCTCGGTGACAACTCCCAAC 345
 |||||
 Db 365 CAGGTCCTCATCTGGCAGGTATCATCCGAGGTGACGCTCGGTGACAACTCCCAAC 424
 |||||
 QY 346 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCTCCGATCATCTCTCTGC 405
 |||||
 Db 425 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCTCCGATCATCTCTCTGC 484
 |||||
 QY 406 CAGGTCAGACACCAAGACAGGGCCACAGAGGAGAGCCAGGTGACGCTGCCCAAGTT 465
 |||||
 Db 485 CAGGTCAGACACCAAGACAGGGCCACAGAGGAGAGCCAGGTGACGCTGCCCAAGTT 544
 |||||
 QY 466 CACCTGCGGAGGAGGCCAGAGAGCTACCGCTCATCGGGTGCCTGCTCTCATATTGC 525
 |||||
 Db 545 CACCTGCGGAGGAGGCCAGAGAGCTACCGCTCATCGGGTGCCTGCTCTCATATTGC 604
 |||||
 QY 526 CTTGTGTGTTTCGCTCATCTCTCTTCAGTTCTG 561
 |||||
 Db 605 CTTGTGTGTTTCGCTCATCTCTCTTCAGTTCTG 640
 |||||

RESULT 15
 BP359643 579 bp mRNA linear EST 17-SEP-2004
 LOCUS BP359643 Sugano cDNA library, mammary gland T47D Homo sapiens cDNA
 DEFINITION BP359643.TDR09260, mRNA sequence.
 ACCESSION BP359643
 VERSION BP359643.1 GI:52289656
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 PUBMED 15342556
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: ysuzuki@ngc.jp.
 Location/Qualifiers
 1..579

FEATURES
 source 1..579

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 07:36:49 ; Search time 1148 Seconds
(without alignments)
10616.281 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 1748
Sequence: 1 ctcagcagaccatcgagagag99.....ggtctctgctcagcagaa 1748

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_8:*

- 1: geneseqn1980s:*
- 2: geneseqn1980s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*
- 15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 1748 | 100.0 | 1748 | 6 | AAD27734 Human tra |
| 2 | 1689 | 96.6 | 1707 | 12 | ADH17439 Human NOV |
| 3 | 1689 | 96.6 | 1708 | 12 | ADH17457 Human NOV |
| 4 | 1687.4 | 96.5 | 1689 | 4 | AAD05796 Human tra |
| 5 | 1687.4 | 96.5 | 1689 | 10 | ADH10392 Human cel |
| 6 | 1687.4 | 96.5 | 1689 | 12 | ADH146916 Human tra |
| 7 | 1687.4 | 96.5 | 1689 | 14 | ADH150145 Human end |
| 8 | 1670.4 | 95.6 | 2380 | 14 | ADZ75551 Human mos |
| 9 | 1670.4 | 95.6 | 2393 | 10 | ADH31742 Human 291 |
| 10 | 1663.4 | 95.2 | 2067 | 4 | AAD05797 Human tra |
| 11 | 1663.4 | 95.2 | 2067 | 10 | ADH10394 Human cel |
| 12 | 1663.4 | 95.2 | 2067 | 12 | ADH146918 Human tra |
| 13 | 1663.4 | 95.2 | 2067 | 14 | ADH150053 Human end |
| 14 | 1645.4 | 94.1 | 2432 | 12 | ADH17413 Human NOV |
| 15 | 1643.8 | 94.0 | 2432 | 12 | ADH17453 Human NOV |
| 16 | 1643.8 | 94.0 | 2432 | 12 | ADH17451 Human NOV |
| 17 | 1643.8 | 94.0 | 2432 | 12 | ADH17449 Human NOV |
| 18 | 1640 | 93.8 | 1771 | 12 | ADH17433 Human NOV |

| | | | | | | |
|----|--------|------|------|----|----------|--------------------|
| 19 | 1640 | 93.8 | 1779 | 12 | ADH17429 | Adh17429 Human NOV |
| 20 | 1479 | 84.6 | 2310 | 12 | ADH17437 | Adh17437 Human NOV |
| 21 | 1429 | 81.8 | 1614 | 6 | ABK31788 | ABK31788 DNA encod |
| 22 | 1402 | 80.2 | 1671 | 14 | AEA19673 | Aea19673 Novel hum |
| 23 | 1369 | 78.3 | 1626 | 12 | ADH17431 | Adh17431 Human NOV |
| 24 | 1302.8 | 74.5 | 1314 | 4 | AAFB3971 | AafB3971 Nucleotid |
| 25 | 1302.8 | 74.5 | 1314 | 4 | AAFB3971 | AafB3971 Human SER |
| 26 | 1302.8 | 74.5 | 1314 | 12 | ADH17441 | Adh17441 Human NOV |
| 27 | 1299.6 | 74.3 | 1314 | 10 | ADE29369 | Ade29369 Human ser |
| 28 | 1278.2 | 73.1 | 1314 | 10 | ADE29366 | Ade29366 Human CDN |
| 29 | 1277.8 | 73.1 | 1574 | 6 | AAD30575 | Aad30575 Human pro |
| 30 | 1258.6 | 72.0 | 1434 | 8 | ABZ22864 | Abz22864 Human den |
| 31 | 1252.8 | 71.7 | 2192 | 4 | AAH14850 | Aah14850 Human CDN |
| 32 | 1155.2 | 66.1 | 1409 | 13 | ADE79066 | Ade79066 Human pro |
| 33 | 1147 | 65.6 | 3103 | 13 | ACN42023 | Acn42023 Human dia |
| 34 | 1143.4 | 65.4 | 1146 | 12 | ADH17443 | Adh17443 Human NOV |
| 35 | 1131.8 | 64.7 | 1407 | 10 | ADE79039 | Ade79039 Human pro |
| 36 | 1120.4 | 64.1 | 1218 | 12 | ADH17435 | Adh17435 Human NOV |
| 37 | 1120.4 | 64.1 | 1218 | 12 | ADH17447 | Adh17447 Human NOV |
| 38 | 1120 | 64.1 | 1448 | 10 | ADE79040 | Ade79040 Human pro |
| 39 | 1105 | 63.2 | 1203 | 12 | ADH17421 | Adh17421 Human NOV |
| 40 | 1103.4 | 63.1 | 1203 | 12 | ADH17427 | Adh17427 Human NOV |
| 41 | 1101.8 | 63.0 | 1203 | 12 | ADH17425 | Adh17425 Human NOV |
| 42 | 1068.4 | 61.1 | 1078 | 4 | AAFB3972 | AafB3972 Nucleotid |
| 43 | 1062.6 | 60.8 | 1341 | 8 | ABZ22865 | Abz22865 Human den |
| 44 | 1053 | 60.2 | 1230 | 6 | AAD27743 | Aad27743 Human tra |
| 45 | 1050 | 60.1 | 1077 | 10 | ADE29368 | Ade29368 Human ser |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAD27734 | standard; DNA; 1748 BP. |
| ID | AAD27734 |
| XX | |
| AC | AAD27734; |
| XX | |
| DT | 18-APR-2002 (first entry) |
| XX | |
| DE | Human transmembrane serine protease DNA. |
| XX | |
| KW | Human; transmembrane serine protease; gene therapy; metastasis; tumour; |
| KW | chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation; |
| KW | atherosclerosis; neurodegenerative disease; neuroprotective; cytoskeletal; |
| KW | pathogenic infection; antiinflammatory; antiarteriosclerotic; |
| KW | antibacterial; ds. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FT | Key |
| FT | Location/Qualifiers |
| FT | 11..1699 |
| FT | /*tag= a |
| FT | /product= "Human transmembrane serine protease" |
| XX | |
| FN | WO200196538-A2. |
| XX | |
| PD | 20-DEC-2001. |
| XX | |
| PF | 12-JUN-2001; 2001WO-EP006618. |
| XX | |
| PR | 13-JUN-2000; 2000US-0211234P. |
| PR | 13-APR-2001; 2001US-0283353P. |
| PR | 16-APR-2001; 2001US-0283648P. |
| XX | |
| PA | (FARB) BAYER AG. |
| XX | |
| PI | Xiao Y, Gedlich R; |
| DR | WPI; 2002-098065/13. |
| DR | P-PSDB; AAE17238. |
| XX | |
| PT | Novel isolated polynucleotide encoding transmembrane serine protease |

PT polypeptide, for treating chronic obstructive pulmonary disease, tumor
PT angiogenesis, inflammation, atherosclerosis and neurodegenerative
PT disease.

PS Claim 1; Page 111-112; 120pp; English.

XX
CC The present invention relates to an isolated polynucleotide encoding a
CC transmembrane serine protease polypeptide. Transmembrane serine protease
CC gene is useful in gene therapy. The invention also relates to a
CC pharmaceutical composition which is useful for modulating the activity of
CC transmembrane serine protease in a disease, such as chronic obstructive
CC pulmonary disease (COPD), metastasis of malignant cells, tumor
CC angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease
CC or pathogenic infection. Transmembrane serine protease is useful as a
CC bait protein in a two-hybrid or three-hybrid assay. The polypeptide is
CC useful for generating antibodies against it and in various assay systems.
CC The present sequence is a human transmembrane serine protease DNA

SQ Sequence 1748 BP; 379 A; 550 C; 487 G; 332 T; 0 U; 0 Other;

Query Match 100.0%; Score 1748; DB 6; Length 1748;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCAGAGACATGAGAGAGACAGCCAGGAAATGATCTTCAGCAAGAACCTTCAGC 60
DB 1 CTCAGAGACATGAGAGAGAGACAGCCAGGAAATGATCTTCAGCAAGAACCTTCAGC 60
QY 61 TGGAGCATCTCCAGCCCGAGGATCTCCAGCTGGGACACCTTCAGGCGGGGATCTCCAGC 120
DB 61 TGGAGCATCTCCAGCCCGAGGATCTCCAGCTGGGACACCTTCAGGCGGGGATCTCCAGC 120
QY 121 CCAGGACATCTCCAGCCCGAGGATCTCCAGCTGGGACACCTTCAGGCGGGGATCTCCAGC 180
DB 121 CCAGGACATCTCCAGCCCGAGGATCTCCAGCTGGGACACCTTCAGGCGGGGATCTCCAGC 180
QY 181 CCAGGACATCTCCAGCTGGTATACCTTCAGGCGGGGATCTTCAGGCGGGGATCTCCAGC 240
DB 181 CCAGGACATCTCCAGCTGGTATACCTTCAGGCGGGGATCTTCAGGCGGGGATCTCCAGC 240
QY 241 CCAGGACATCTCCAGCTGGTATACCTTCAGGCGGGGATCTTCAGGCGGGGATCTCCAGC 300
DB 241 CCAGGACATCTCCAGCTGGTATACCTTCAGGCGGGGATCTTCAGGCGGGGATCTCCAGC 300
QY 301 CAGGTCATCATCCGCCAGGTCAGGCTCGGTGACAACTCCCAACAGAGTGTAACCTTGT 360
DB 301 CAGGTCATCATCCGCCAGGTCAGGCTCGGTGACAACTCCCAACAGAGTGTAACCTTGT 360
QY 361 TAGAGCAACACAGTGGGGGCTGTACCATCGATCATCTCTCGCAGAGTCAAGCAGC 420
DB 361 TAGAGCAACACAGTGGGGGCTGTACCATCGATCATCTCTCGCAGAGTCAAGCAGC 420
QY 421 AACCAAGGCGCCAGGAGAGACCCAGATGACAGGCTGCGCAAGTTACCTTGCGGAGG 480
DB 421 AACCAAGGCGCCAGGAGAGACCCAGATGACAGGCTGCGCAAGTTACCTTGCGGAGG 480
QY 481 CCAGAAACAGGTACCGCTCATCGGAGGTGCTCTCTCATTTGCCCTGAGTTTGGCT 540
DB 481 CCAGAAACAGGTACCGCTCATCGGAGGTGCTCTCTCATTTGCCCTGAGTTTGGCT 540
QY 541 CATCATCTCTTTCAGTTTGGCAGGGGCAACAAGGATCAAGTAAAGAGAGAGGA 600
DB 541 CATCATCTCTTTCAGTTTGGCAGGGGCAACAAGGATCAAGTAAAGAGAGAGGA 600
QY 601 GAGCTGTCCCAAGAGAGCTGTGTGCTGTGACGGGGGTGTGAGCTGCAAGCTGAAAGAGGA 660
DB 601 GAGCTGTCCCAAGAGAGCTGTGTGCTGTGACGGGGGTGTGAGCTGCAAGCTGAAAGAGGA 660
QY 661 CGAGCTGGGCTGCTGAGGTTTGAATGAGCAAGTCTCTGTTAAATCTACTCTGGGTC 720
DB 661 CGAGCTGGGCTGCTGAGGTTTGAATGAGCAAGTCTCTGTTAAATCTACTCTGGGTC 720
QY 721 CTCCATCATGAGCTTCCCATCTGTAGACAGCACTGAATGACTCTTACTCAGAGAAAGAC 780

DB 721 CTCCATCATGAGCTTCCCATCTGTAGACAGCACTGAATGACTCTTACTCAGAGAAAGAC 780
QY 781 CTGCGACAGCTGGGCTTTCAGAGGTGCTCAACCGGACAAACGAGGTTGCCACAGGATTT 840
DB 781 CTGCGACAGCTGGGCTTTCAGAGGTGCTCAACCGGACAAACGAGGTTGCCACAGGATTT 840
QY 841 TGCCAAAGCTTTCATCTTGAATATCAATCCAGCATCCAGGAAAGGCTCCACAGGTC 900
DB 841 TGCCAAAGCTTTCATCTTGAATATCAATCCAGCATCCAGGAAAGGCTCCACAGGTC 900
QY 901 TGAATGCTTTCAGGCTATATCTCTCCAGTGTTCCTCCAGCTGCGAGCTGAGGCGCAT 960
DB 901 TGAATGCTTTCAGGCTATATCTCTCCAGTGTTCCTCCAGCTGCGAGCTGAGGCGCAT 960
QY 961 GACCGGCGGATCTGGAGAGGGGCGCTGGGCTGGGATGACAAAGTGGCTTGGCAAGTGAG 1020
DB 961 GACCGGCGGATCTGGAGAGGGGCGCTGGGCTGGGATGACAAAGTGGCTTGGCAAGTGAG 1020
QY 1021 TCTGACCTTGGGACCAACCAATCTGTGAGAGCAGCTCATTTGACCGCCAGTGGTGTCT 1080
DB 1021 TCTGACCTTGGGACCAACCAATCTGTGAGAGCAGCTCATTTGACCGCCAGTGGTGTCT 1080
QY 1081 CACTGCGGCCCATCTGCTTCTGCTGACCCGAGAGAAAGTCTTGGAGGGCTGGAAGGTGTA 1140
DB 1081 CACTGCGGCCCATCTGCTTCTGCTGACCCGAGAGAAAGTCTTGGAGGGCTGGAAGGTGTA 1140
QY 1141 CGCGGGGACAGCAACCTGACCCAGTGGCTGAGGAGGAGGCTCCATTCGCGAGATCATCAT 1200
DB 1141 CGCGGGGACAGCAACCTGACCCAGTGGCTGAGGAGGAGGCTCCATTCGCGAGATCATCAT 1200
QY 1201 CAACAGCAATTACACCGATGAGAGAGAGCACTATGATGATGCGCTCATGCGGCTGTCCAA 1260
DB 1201 CAACAGCAATTACACCGATGAGAGAGAGCACTATGATGATGCGCTCATGCGGCTGTCCAA 1260
QY 1261 GCGCTTGAACCTTGTGCTGCTCATCAACCTGCTGCTTCCCATGCAATGAGACAGACTT 1320
DB 1261 GCGCTTGAACCTTGTGCTGCTCATCAACCTGCTGCTTCCCATGCAATGAGACAGACTT 1320
QY 1321 TAGGCTCAATGAGACCTGCTGATCAAGGCTTGGCAAGGACCAAGGAGACAGATGACAA 1380
DB 1321 TAGGCTCAATGAGACCTGCTGATCAAGGCTTGGCAAGGACCAAGGAGACAGATGACAA 1380
QY 1381 GACATCCGCCCTCTCCCGGAGGTGACAGTCAATCTCATGATGATCAAGAAATGCAATGA 1440
DB 1381 GACATCCGCCCTCTCCCGGAGGTGACAGTCAATCTCATGATGATCAAGAAATGCAATGA 1440
QY 1441 CTACTTGGCTTATGACAGTTAATCTTACCCCAAGGATGATGTGCTGGGACCTTGTGTG 1500
DB 1441 CTACTTGGCTTATGACAGTTAATCTTACCCCAAGGATGATGTGCTGGGACCTTGTGTG 1500
QY 1501 GGGGAGAGACTCCGCGAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACACG 1560
DB 1501 GGGGAGAGACTCCGCGAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACACG 1560
QY 1561 CTGGTACCTTGGCAGAGTGTACACAGCTGGGACAGGCTGTGTGGCCAGAGAAACAACTGG 1620
DB 1561 CTGGTACCTTGGCAGAGTGTACACAGCTGGGACAGGCTGTGTGGCCAGAGAAACAACTGG 1620
QY 1621 TGTGTACACCAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGACGAGGT 1680
DB 1621 TGTGTACACCAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGACGAGGT 1680
QY 1681 GCGATTCAGAAATCTTAACAGCTGGCTGCTGTGACACAGACACCGGCTGCTGTGAC 1740
DB 1681 GCGATTCAGAAATCTTAACAGCTGGCTGCTGTGACACAGACACCGGCTGCTGTGAC 1740
QY 1741 TCGAGAAA 1748
DB 1741 TCGAGAAA 1748

RESULT 2


```

Db 670 GGCTGCGTAGGTTGACTGGGACAAGCTCTGCTTAAATCTACTCTGGGCTCTCCCAT
Qy 728 CAGTGCTTCCATCTGTAGAGCACTGGAAATGATCTCTAATGAGAAAGACTTGCAG
Db 730 CAGTGCTTCCATCTGTAGAGCACTGGAAATGATCTCTAATGAGAAAGACTTGCAG
Qy 788 CAGTGCTTCCATCTGTAGAGCACTGGAAATGATCTCTAATGAGAAAGACTTGCAG
Db 790 CAGTGCTTCCATCTGTAGAGCACTGGAAATGATCTCTAATGAGAAAGACTTGCAG
Qy 848 AGCTTCTCATCTTGAATACAACTCCACCTCAGAAAGCTTCCACAGTCTGAATGC
Db 850 AGCTTCTCATCTTGAATACAACTCCACCTCAGAAAGCTTCCACAGTCTGAATGC
Qy 908 CTTTCCAGAGGGGTATCTCTCCCTCAGGTTCCACCTGCGGACGAGGGCCATGACCGGG
Db 910 CTTTCCAGAGGGGTATCTCTCCCTCAGGTTCCACCTGCGGACGAGGGCCATGACCGGG
Qy 968 CGGATCGTGGAGGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAATCTGCAC
Db 970 CGGATCGTGGAGGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAATCTGCAC
Qy 1028 TTGCGGACACCCACATCTGTGAGGCAAGCTCATTTGACGCCCAAGTGGTCTCACTGCC
Db 1030 TTGCGGACACCCACATCTGTGAGGCAAGCTCATTTGACGCCCAAGTGGTCTCACTGCC
Qy 1088 GCCCACTGCTTCTTGTGAGACCCGGGAGAAAGTCTGGAGGGCTGGAAGGTATAGCCGGGC
Db 1090 GCCCACTGCTTCTTGTGAGACCCGGGAGAAAGTCTGGAGGGCTGGAAGGTATAGCCGGGC
Qy 1148 ACCAGCAACCTGACCAAGTTGCTGAGGCAAGCTCATTTGCGAGATCATCATCAAGC
Db 1150 ACCAGCAACCTGACCAAGTTGCTGAGGCAAGCTCATTTGCGAGATCATCATCAAGC
Qy 1208 AATTACACCGATGAGAGGAGATGATGACATGCGCTCATGCGGCTGTCCAAAGCCCTG
Db 1210 AATTACACCGATGAGAGGAGATGATGACATGCGCTCATGCGGCTGTCCAAAGCCCTG
Qy 1268 ACCGTGCGGCTCATGACACCTGCTTGCCTCCCATGATGAGACCTTTAGGCTC
Db 1270 ACCGTGCGGCTCATGACACCTGCTTGCCTCCCATGATGAGACCTTTAGGCTC
Qy 1328 AATGAGACCTGCTGATCATCAGGCTTTGGCAAGACCAAGGAGACAGATGACAAAGATCC
Db 1330 AATGAGACCTGCTGATCATCAGGCTTTGGCAAGACCAAGGAGACAGATGACAAAGATCC
Qy 1388 CCGTTCTCCGGAGGCTGAGGCTGATCTCATGCACTTCAAAGAAATGATGACTTGG
Db 1390 CCGTTCTCCGGAGGCTGAGGCTGATCTCATGCACTTCAAAGAAATGATGACTTGG
Qy 1448 GTCTATGACAGTTACTTTACCCCAAGATGATGATGCTGGGAGCTTCCGAGGGGGAGA
Db 1450 GTCTATGACAGTTACTTTACCCCAAGATGATGATGCTGGGAGCTTCCGAGGGGGAGA
Qy 1508 GACTCTGCGCAGGAGACAGCGGGGGGCTCTTGTCTGTGACGAGAAACAACCGTGTATC
Db 1510 GACTCTGCGCAGGAGACAGCGGGGGGCTCTTGTCTGTGACGAGAAACAACCGTGTATC
Qy 1568 CTGGCAGGTGTACACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTGTAC
Db 1570 CTGGCAGGTGTACACAGCTGGGGGACAGGCTGTGTGGCCAGAGAAACAAACCTGTGTGTAC
Qy 1628 ACCAAGTGAAGAAGTTCTTCCCGATTATACAGCAAGATGAGAGGAGAGTGCATTC
Db 1630 ACCAAGTGAAGAAGTTCTTCCCGATTATACAGCAAGATGAGAGGAGAGTGCATTC
Qy 1688 AGAAATCC 1696
Db 1690 AGAAATCC 1698

```

RESULT 3

```

ADH17457/C
ID ADH17457 standard; cDNA; 1708 BP.
XX
AC ADH17457;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human NOV13b cDNA - SEQ ID 147.
XX
KW NOX; antidiabetic; anorectic; cardiant; hypotensive;
KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW protozoacide; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW dermatological; antiaesthetic; antilipemic; metabolic; diabetes;
KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
KW epilepsy; immune; osteoarthritis; haemopoietic;
KW inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
KW cell differentiation; proliferation; haemopoiesis; wound healing;
KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
KW pharmacogenomic; human; ss; gene.
OS Homo sapiens.
XX
PN WO2003093432-A2.
XX
PD 13-NOV-2003.
XX
PE 02-MAY-2003; 2003WO-US013690.
XX
PR 02-MAY-2002; 2002US-0377321P.
PR 08-MAY-2002; 2002US-0378730P.
PR 24-MAY-2002; 2002US-0383075P.
PR 29-MAY-2002; 2002US-0384044P.
PR 30-MAY-2002; 2002US-0384215P.
PR 30-MAY-2002; 2002US-0384236P.
PR 30-MAY-2002; 2002US-0384297P.
PR 30-MAY-2002; 2002US-0384337P.
PR 30-MAY-2002; 2002US-0384352P.
PR 31-MAY-2002; 2002US-0385211P.
PR 02-JUL-2002; 2002US-0393333P.
PR 09-AUG-2002; 2002US-0402154P.
PR 09-AUG-2002; 2002US-0402171P.
PR 09-AUG-2002; 2002US-0402204P.
PR 09-AUG-2002; 2002US-0402205P.
PR 09-AUG-2002; 2002US-0405175P.
PR 22-AUG-2002; 2002US-0406129P.
PR 27-AUG-2002; 2002US-0412954P.
PR 23-SEP-2002; 2002US-0412954P.
PR 30-SEP-2002; 2002US-0414975P.
PR 07-OCT-2002; 2002US-0416611P.
PR 24-OCT-2002; 2002US-0420851P.
PR 31-OCT-2002; 2002US-0422547P.
PR 01-MAY-2003; 2003US-00428275.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alvarez E, Anderson DW, Boldog FL, Catterton E, Edinger SR,
PI Fernandes ER, Gerlach VU, Gorman L, Grose WM, Guo X, Ji W,
PI Kehuda R, Li L, Macdougall JR, Padigaru M, Patturajan M,
PI Verneton JD, Rastelli L, Shinkets RA, Szytek KA, Stone DJ,
PI Vernet CM, Voss EZ, Zhong M;
XX
WP; 2004-053040/05.
DR P-PSDB; ADH17458.
XX
PT New isolated NOX polypeptide, useful for preventing, diagnosing or
PT treating NOX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
XX
PS Claim 20; SEQ ID NO 147; 478pp; English.
XX
CC The invention relates to a novel isolated NOX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,

```


CC cardiant, hypotensive, antiarteriosclerotic, anorectic, vinuclide,
CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic,
CC antiinflammatory, dermatological, antisthmatic and antidiabetic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, aschma and various types of dyslipidemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX cDNA of the invention.

XX Sequence 1708 BP; 323 A; 477 C; 539 G; 369 T; 0 U; 0 Other;

Query Match 96.6%; Score 1689; DB 12; Length 1708;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ACCATGAGAGGAGGACAGCCAGGGATGCACTTCACAGAAACACCTTCAGCTGAGCA 67
DB 1698 ACCATGAGAGGAGGACAGCCAGGGATGCACTTCACAGAAACACCTTCAGCTGAGCA 1639
QY 68 TCTCCAGCCCAAGCATCTCCAGCTGGAGACCTTCAGCGCGGATCTCCAGCCAGCA 127
DB 1638 TCTCCAGCCCAAGCATCTCCAGCTGGAGACCTTCAGCGCGGATCTCCAGCCAGCA 1579
QY 128 TCTCCAGCCCAAGCATCTCCAGCTGGAGACCTTCAGCGCGGATCTCCAGCCAGCA 187
DB 1578 TCTCCAGCCCAAGCATCTCCAGCTGGAGACCTTCAGCGCGGATCTCCAGCCAGCA 1519
QY 188 TCTCCAGCTGGTACACCTCCAGCGCGGATCTTCAGAGCGGATCTCCAGCCAGCA 247
DB 1518 TCTCCAGCTGGTACACCTCCAGCGCGGATCTTCAGAGCGGATCTCCAGCCAGCA 1459
QY 248 TCTCCAGCGCGGATCTCCAGCTGGAGACCTTCAGAGCTTCAGCTGGAGGATCA 307
DB 1458 TCTCCAGCGCGGATCTCCAGCTGGAGACCTTCAGAGCTTCAGCTGGAGGATCA 1399
QY 308 TCTCCAGCGGATCTCCAGCTGGAGACCTTCAGAGCTTCAGCTGGAGGATCA 367
DB 1398 TCTCCAGCGGATCTCCAGCTGGAGACCTTCAGAGCTTCAGCTGGAGGATCA 1339
QY 368 ACACCAAGTGGGGGCTGTACCATCCGATCATCTCTGCAAGTACAGCAACAGG 427
DB 1338 ACACCAAGTGGGGGCTGTACCATCCGATCATCTCTGCAAGTACAGCAACAGG 1279
QY 428 GCCACCAAGGAGAGCCAGGATCAGAGCTTCGCAAGTTACCTGCGGAGAGGCGCAAG 487
DB 1278 GCCACCAAGGAGAGCCAGGATCAGAGCTTCGCAAGTTACCTGCGGAGAGGCGCAAG 1219
QY 488 CAGTACCGCTCATCGGGTGGGTCTCTCTCATTTGCCCTGGGTGTTTGGCTCATATC 547
DB 1218 CAGTACCGCTCATCGGGTGGGTCTCTCTCATTTGCCCTGGGTGTTTGGCTCATATC 1159
QY 548 CTCTTCAAGTTCTGGAGGCGCACAGAGATCAGGTACAGAGAGAGGAGAGGCTGT 607
DB 1158 CTCTTCAAGTTCTGGAGGCGCACAGAGATCAGGTACAGAGAGAGGAGAGGCTGT 1039
QY 608 CCCAAGGACGCTGTTCCGTGTGACGGGGTGGTGACTGCAAGCTGAAGAGTGAAGAGCTG 667
DB 1098 CCCAAGGACGCTGTTCCGTGTGACGGGGTGGTGACTGCAAGCTGAAGAGTGAAGAGCTG 1039
QY 668 GGCTGCGGAGGTTGACTGGGACAGTCTGTGCTAAATCTACTGGGCTCTCCAT 727

DB 1038 GGCTGCGGAGGTTGACTGGGACAGTCTGTGCTAAATCTACTGGGCTCTCCAT 979
QY 728 CAGTGGCTTCCCATCTGTACAGCACTGGAATGACTCTTACTCAGAGAGACTGCGAC 787
DB 978 CAGTGGCTTCCCATCTGTACAGCACTGGAATGACTCTTACTCAGAGAGACTGCGAC 919
QY 788 CAGTGGGTTTCCAGAGTGTCAACCGGACCAACCGAGGTTGCCACAGGATTTGGCAAC 847
DB 918 CAGTGGGTTTCCAGAGTGTCAACCGGACCAACCGAGGTTGCCACAGGATTTGGCAAC 859
QY 848 AGCTTCTCAATCTTGAGATCAACTCCACCATCCAGGAAAGCTTCCACAGCTGAATGC 907
DB 858 AGCTTCTCAATCTTGAGATCAACTCCACCATCCAGGAAAGCTTCCACAGCTGAATGC 799
QY 908 CTTTCCAGCGGTATATCTCCCTTCAGTGTTCCTCACTGCGGACTGAGGCGCATGACCGG 967
DB 798 CTTTCCAGCGGTATATCTCCCTTCAGTGTTCCTCACTGCGGACTGAGGCGCATGACCGG 739
QY 968 CGGATCGTGGAGGAGGCGGCTGGGCTGGATAGCAAGTGGGCTTGGCAAGTGAATCTGCAC 1027
DB 738 CGGATCGTGGAGGAGGCGGCTGGGCTGGATAGCAAGTGGGCTTGGCAAGTGAATCTGCAC 679
QY 1028 TTGGGACACACCAACATCTGTGAGAGGACGCTGATTTGACGCCAGTGCTTCACTGCC 1087
DB 678 TTGGGACACACCAACATCTGTGAGAGGACGCTGATTTGACGCCAGTGCTTCACTGCC 619
QY 1088 GCCCATGCTTCTTCTGTCACCCGGAGAGAGTCTGGAGGCTGGAAGGTGACCGGGC 1147
DB 618 GCCCATGCTTCTTCTGTCACCCGGAGAGAGTCTGGAGGCTGGAAGGTGACCGGGC 559
QY 1148 ACCAGCAACCTGACCAAGTTCCTGAGGAGGCTCCATTCGCGAGATCATCAACAGC 1207
DB 558 ACCAGCAACCTGACCAAGTTCCTGAGGAGGCTCCATTCGCGAGATCATCAACAGC 499
QY 1208 AATTACACCATGAGAGAGAGCACTATGATGACCTGCTTCAAGGCTGTCGAAGCCCTG 1267
DB 498 AATTACACCATGAGAGAGAGCACTATGATGACCTGCTTCAAGGCTGTCGAAGCCCTG 439
QY 1268 ACCCTGTGCTGCATCATCCCTGCTTGCCTCCCATGATGAGACAGACCTTTAGCCTC 1327
DB 438 ACCCTGTGCTGCATCATCCCTGCTTGCCTCCCATGATGAGACAGACCTTTAGCCTC 379
QY 1328 AATGAGACTGCTGTGATCAAGGCTTTGGCAAGCCAGGAGAGAGATGACAAAGATCC 1387
DB 378 AATGAGACTGCTGTGATCAAGGCTTTGGCAAGCCAGGAGAGAGATGACAAAGATCC 319
QY 1388 CCGTCTCCGGAGGATGAGGATCAATCTCATGCACTTCAAGAAATGACATGACTATG 1447
DB 318 CCGTCTCCGGAGGATGAGGATCAATCTCATGCACTTCAAGAAATGACATGACTATG 259
QY 1448 GTCTATGACAGTTACCTTACCCCAAGGATGATGTGTCGGGAGCTTCTGCGGGGAG 1507
DB 258 GTCTATGACAGTTACCTTACCCCAAGGATGATGTGTCGGGAGCTTCTGCGGGGAG 199
QY 1508 GACTCTGCGAGGAGAGCAAGCGGGGCTCTTGTCTGTGACGAGAACACCGCTGTGAC 1567
DB 198 GACTCTGCGAGGAGAGCAAGCGGGGCTCTTGTCTGTGACGAGAACACCGCTGTGAC 139
QY 1568 CTGGCAGGTGTCACAGCTGGGAGCAGGCTGTGAGCAGAGAAACCAACCTGTGTATC 1627
DB 138 CTGGCAGGTGTCACAGCTGGGAGCAGGCTGTGAGCAGAGAAACCAACCTGTGTATC 79
QY 1628 ACCAAGTGAAGAAGTTCTTCTCGATTTTACAGCAAGATGAGAGAGAGTGCATTC 1687
DB 78 ACCAAGTGAAGAAGTTCTTCTCGATTTTACAGCAAGATGAGAGAGAGTGCATTC 19
QY 1688 AGAAAATCC 1696
DB 18 AGAAAATCC 10

RESULT 4

AAD05796
 ID AAD05796 standard; DNA; 1689 BP.
 XX
 AC AAD05796;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Human transmembrane serine protease (Endothelinase 2-S) DNA.
 XX
 KM Human: endothelinase 2-S; protease domain; cytosolic; vulnery; wound;
 KM angiotensin; peridontitis; dermatological disorder; gene therapy; scar;
 KM angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
 KM chronic inflammatory disease; ocular disorder; circulatory disorder;
 KM crest syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
 KM liver cirrhosis; osteoradionecrosis; systemic sclerosis; oesophageal;
 KM inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
 KM systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
 KM transmembrane serine protease; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT 1..1689
 FT CDS /tag= a
 FT /product= "Human endothelinase 2-S protein"
 XX
 PN M0200136604-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000MO-US031803.
 XX
 PR 18-NOV-1999; 99US-0166391P.
 PR 22-SEP-2000; 2000US-0234840P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Madison BL, Ong EO;
 XX
 DR WPI: 2001-336001/35.
 DR P-PSDB; AAE01943.
 XX
 PT New nucleic acid encoding a protein comprising endothelinase activly
 PT useful in the prevention and treatment of e.g. vascular malformations,
 PT cardiovascular disorders, and chronic inflammatory disease.
 XX
 PS Example 2; Page 135-137; 152pp; English.
 CC The present DNA sequence encodes human short form transmembrane serine
 CC protease (Endothelinase 2-S) protein. The invention relates to an
 CC endothelinase protein, endothelinase protease domain and their
 CC corresponding nucleic acid molecules. An endothelinase protein or protease
 CC domain of it is useful for the treatment and diagnosis of disorders
 CC associated with aberrant angiogenesis or undesired neovascularisation.
 CC The undesired angiogenesis is associated with disorders selected from
 CC solid neoplasm, vascular malformations and cardiovascular disorders such
 CC as angiodioma, angiolipoma, atherosclerosis, restenosis/reperfusion
 CC injury, arteriovenous malformations, haemangiomas and vascular
 CC adhesions, dyschondroplasia with vascular hamartomas (Pafucis
 CC syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
 CC syndrome) and Von Hippel Lindau syndrome, chronic inflammatory bowel
 CC disease, nonhealing fractures, haemophilic joints, psoriasis, rheumatoid
 CC arthritis, venous stasis ulcers, granulations-burns, hypertrophic scars,
 CC liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
 CC granuloma and systemic sclerosis and aberrant wound repair, circulatory
 CC disorders Raynaud's phenomenon, crest syndrome such as calcinosis,
 CC oesophageal, dyomolity, sclerodactyly and teangiectasis, dermatological
 CC disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
 CC vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine
 CC stains, blue rubber blab nevus syndrome, Klippel-Trenaunay-Weber syndrome
 CC and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
 CC caused by ocular neovascular disease, corneal graft neovascularisation,

CC macular degeneration, retinopathy of prematurity, retrolental fibroplasia
 CC and corneal neovascularisation. The nucleic acids of the invention are
 CC also used in gene therapy. The invention also provides method for
 CC screening compounds that modulate angiogenesis
 XX
 SO Sequence 1689 BP; 367 A; 529 C; 470 G; 323 T; 0 U; 0 Other;
 Query Match 96.5%; Score 1687.4; DB 4; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 11 ATGAGAGGACACACCGGGAATGCATCTCCAGACAGAAACCTTCAGCTGGAGCATCT 70
 1 ATGAGAGGACACACCGGGAATGCATCTCCAGACAGAAACCTTCAGCTGGAGCATCT 60
 71 CCAGCCGAGCATCTCCAGCTGGAGACACCTCCAGGCGGAGCATCTCCAGCCGAGCATCT 130
 61 CCAGCCGAGCATCTCCAGCTGGAGACACCTCCAGGCGGAGCATCTCCAGCCGAGCATCT 120
 131 CCAGCCGAGCATCTCCAGCTGGAGACACCTCCGAGCGGAGCATCTCCAGCCGAGCATCT 190
 121 CCAGCCGAGCATCTCCAGCTGGAGACACCTCCGAGCGGAGCATCTCCAGCCGAGCATCT 180
 191 CCAGCTGGATACCTCCAGGCGGAGCATCTCCAGGCGGAGCATCTCCAGCCGAGCATCT 250
 181 CCAGCTGGATACCTCCAGGCGGAGCATCTCCAGGCGGAGCATCTCCAGCCGAGCATCT 240
 251 CCAGCCGAGCATCTCCAGCTGGAGACACCTCCAGCTGGATCTCCAGCTGGAGCATCTCA 310
 241 CCAGCCGAGCATCTCCAGCTGGAGACACCTCCAGCTGGATCTCCAGCTGGAGCATCTCA 300
 301 TCGGCGAGGATCAGGCTCGGTGACAACTCCCAACAGAGTACTCTTTAGAGCAACA 360
 371 CCAATGGGGGCTGTACCATTCCTATCTCTCTGAGTCAAGACCAAGCAACAGGGCC 430
 361 CCAATGGGGGCTGTACCATTCCTATCTCTCTGAGTCAAGACCAAGCAACAGGGCC 420
 431 ACCAGGAGAGCCAGGATACAGGCTGCCCAAGTTCACCTGAGGAGGAGGAGCAAGCAG 490
 421 ACCAGGAGAGCCAGGATACAGGCTGCCCAAGTTCACCTGAGGAGGAGGAGCAAGCAG 480
 491 CTACCGCTCATCGGGTGCCT 550
 481 CTACCGCTCATCGGGTGCCT 540
 551 TTCCAGTTCTGGACAGGAGCCACAGGATACAGGATACAGGATACAGGATACAGGATAC 610
 541 TTCCAGTTCTGGACAGGAGCCACAGGATACAGGATACAGGATACAGGATACAGGATAC 600
 611 AAGCAGCTGTGCTGTGACGAGGAGTGTGAGTCAAGCTGAAAGTGAAGAGCTGGGC 670
 601 AAGCAGCTGTGCTGTGACGAGGAGTGTGAGTCAAGCTGAAAGTGAAGAGCTGGGC 660
 671 TGCCTGAGGTTTGACTGGGACAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 730
 661 TGCCTGAGGTTTGACTGGGACAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 731 TGGCTTCCCATCTGTACAGCAACTGGAATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 790
 721 TGGCTTCCCATCTGTACAGCAACTGGAATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 791 CTGGGTTTTCAGAGTGTCTACCGGACCAAGGAGTGTCTCCCAAGGAGTTTTCAGAGC 850
 781 CTGGGTTTTCAGAGTGTCTACCGGACCAAGGAGTGTCTCCCAAGGAGTTTTCAGAGC 840
 851 TTCTCAATCTTGAGATCAACTCCAGCATCCAGAAAGCTTCCACAGTCTGAATGCTCT 910
 841 TTCTCAATCTTGAGATCAACTCCAGCATCCAGAAAGCTTCCACAGTCTGAATGCTCT 900
 911 TCCAGGAGGATATCT 970

Db 901 TCCACGCGTATATCTCCCTCCAGTGTCCCACTGCGGAGCATGAGGCCATGACCGGGCGG 960
 Qy 971 ATGCTGGAGAGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1030
 Db 961 ATCTGGAGAGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1020
 Qy 1031 GGCACCAACCAATCTGTGAGGAGCGCTCATTAAGACCCCAATGGGAGTCTCACTGCGGCC 1090
 Db 1021 GGCACCAACCAATCTGTGAGGAGCGCTCATTAAGACCCCAATGGGAGTCTCACTGCGGCC 1080
 Qy 1091 CACTGCTTCTTCTGAGCCCGGAGAGAGTCTGAGAGGCTGGAAGGTGACGCGGGACCC 1150
 Db 1081 CACTGCTTCTTCTGAGCCCGGAGAGAGTCTGAGAGGCTGGAAGGTGACGCGGGACCC 1140
 Qy 1151 AGCAACTGTGACCGTGGCTGAGGAGCGCTTCCATTCGCCAGATCATCATTAACAGCAAT 1210
 Db 1141 AGCAACTGTGACCGTGGCTGAGGAGCGCTTCCATTCGCCAGATCATCATTAACAGCAAT 1200
 Qy 1211 TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGACC 1270
 Db 1201 TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGACC 1260
 Qy 1271 CTGTCCGCTCATCATTCACCTCTGTGCTTCCCATGATGAGACACCTTTCAGCTCAAT 1330
 Db 1261 CTGTCCGCTCATCATTCACCTCTGTGCTTCCCATGATGAGACACCTTTCAGCTCAAT 1320
 Qy 1331 GAGACCTGTGATCATCAGCTTTGGCAAGACCGAGAGACATGACAGACATCCCC 1390
 Db 1321 GAGACCTGTGATCATCAGCTTTGGCAAGACCGAGAGACATGACAGACATCCCC 1380
 Qy 1391 TTCTCCGGGGGAGGAGGATCAATCTCATGCACTTCAAGAAATGACATGACTTACCTGCTC 1450
 Db 1381 TTCTCCGGGGGAGGAGGATCAATCTCATGCACTTCAAGAAATGACATGACTTACCTGCTC 1440
 Qy 1451 TATGACATTAATCTTACCTTCCCAAGATGATGTGCTGGGAGACTTCTGTGGGGGAGAGAC 1510
 Db 1441 TATGACATTAATCTTACCTTCCCAAGATGATGTGCTGGGAGACTTCTGTGGGGGAGAGAC 1500
 Qy 1511 TCCTGCCAGGAGAGACAGCGGGGGCTTGTGTGTGAGCAAGAAACACCGTGTGACTGTG 1570
 Db 1501 TCCTGCCAGGAGAGACAGCGGGGGCTTGTGTGTGAGCAAGAAACACCGTGTGACTGTG 1560
 Qy 1571 GCAGGTGTCACCACTGGGGGAGAGCGGTGTGCGAGAGAAACCAACCTGGTGTACACCC 1630
 Db 1561 GCAGGTGTCACCACTGGGGGAGAGCGGTGTGCGAGAGAAACCAACCTGGTGTACACCC 1620
 Qy 1631 AAAGTGAAGAGTCTTCTGATTTTACAGAAATGAGAGCGAGTGCATTCAGA 1690
 Db 1621 AAAGTGAAGAGTCTTCTGATTTTACAGAAATGAGAGCGAGTGCATTCATA 1680
 Qy 1691 AAATCTTAA 1699
 Db 1681 AAATCTTAA 1689
 RESULT 5
 ADI10392 ID ADI10392 standard; DNA; 1689 BP.
 XX AC ADI10392;
 XX 22-APR-2004 (first entry)
 XX Human cell surface protease coding sequence #12.
 XX therapeutic agent; plasmin; protease specific antigen; PSA;
 KW cell-surface protease-associated disease; cancer; ocular disease;
 KW cardiovascular disease; chronic inflammatory disease; wound;
 KW circulatory disorder; dermatological disorder; rheumatoid arthritis;
 KW psoriasis; diabetic retinopathy; pterygium;
 KW excimer laser surgery scarring; glaucoma filtering surgery scarring;
 KW macular degeneration; crest syndrome; solid neoplasm; vascular tumour;
 KW melanoma; Kaposi's sarcoma; human; cell surface protease; gene; ds.

XX OS Homo sapiens.
 XX PN W0200295007-A2.
 XX PD 28-NOV-2002.
 XX PF 23-MAY-2002; 2002W0-US016819.
 XX PR 23-MAY-2001; 2001US-0293267P.
 XX (CORV-) CORVAS INT INC.
 PI Madison EL, Sempke JE, Vlaeuk GP, Kemp SJ, Komandla M, Siev DV;
 PI WPI; 2003-221280/21.
 DR P-PSDB; ADI10393.
 XX Novel conjugate useful for treating cell-surface protease-associated
 PT disease, comprises a therapeutic agent and a peptidic or nucleic acid
 PT substrate linked to it optionally by a peptidic linker.
 PS Claim 9; SEQ ID NO 23; 581bp; English.
 CC The invention comprises a conjugate that consists of a therapeutic agent
 CC and a peptide substrate (optionally linked via linker). The peptide
 CC substrate is proteolytically cleaved by a cell surface protease pr a
 CC soluble, released or shed form of it, to liberate the therapeutic agent,
 CC the conjugate of the invention is not substantially cleaved by plasmin or
 CC protease specific antigen (PSA). The conjugate of the invention is useful
 CC for treating a cell-surface protease-associated disease such as: cancer,
 CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
 CC wounds, circulatory disorders, dermatological disorders, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
 CC scarring from excimer laser surgery, scarring from glaucoma filtering
 CC surgery, macular degeneration, crest syndromes, solid neoplasms, vascular
 CC tumours, melanoma and Kaposi's sarcoma. The present DNA sequence encodes
 CC a human cell surface protease.
 XX Sequence 1689 BP; 367 A; 529 C; 470 G; 323 T; 0 U; 0 Other;
 SQ
 Query Match 96.5%; Score 1687.4; DB 10; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 11 ATGAGAGGAGACAGCGGAGATGATCTTCACAGAAAGAACCTTACGCTGAGCATCT 70
 Db 1 ATGAGAGGAGACAGCGGAGATGATCTTCACAGAAAGAACCTTACGCTGAGCATCT 60
 Qy 71 CCAGCCCAAGCATCTCCAGCTGGGAGACCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 130
 Db 61 CCAGCCCAAGCATCTCCAGCTGGGAGACCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 120
 Qy 131 CCAGCCCAAGCATCTCCAGCTGGGAGACCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 190
 Db 121 CCAGCCCAAGCATCTCCAGCTGGGAGACCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 180
 Qy 191 CCAGCTGTACACCTCCAGGCGGGGATCTTCAGAGCGGGGATCTCCAGCCCAAGCATCT 250
 Db 181 CCAGCTGTACACCTCCAGGCGGGGATCTTCAGAGCGGGGATCTCCAGCCCAAGCATCT 240
 Qy 251 CCAGCCCGGGGATCTCCGGGCTTGGGATCACTTTCAGAGTCTTCCAGGAGGATCTCA 310
 Db 241 CCAGCCCGGGGATCTCCGGGCTTGGGATCACTTTCAGAGTCTTCCAGGAGGATCTCA 300
 Qy 311 TCCGCGAGGTGAGCTGTGAGCAACCTCCCAACAGAGTGAATCTTGTAGAGCAACA 370
 Db 301 TCCGCGAGGTGAGCTGTGAGCAACCTCCCAACAGAGTGAATCTTGTAGAGCAACA 360
 Qy 371 CCAATGGGGGCTGATCCATTCATATCTCTGCAAGTTCAGACCAAGCAACGAGGCC 430
 Db 361 CCAATGGGGGCTGATCCATTCATATCTCTGCAAGTTCAGACCAAGCAACGAGGCC 420

| | | | |
|----|------|--|------|
| QY | 431 | ACGAGGGAGAGACCCAGGATACAGAGCTGGCCAAAGTTACCTGGCGGGAGGGCCAGAACGAG | 490 |
| Db | 421 | ACCGAGGGAGAGACCCAGGATACAGAGCTGGCCAAAGTTACCTGGCGGGAGGGCCAGAACGAG | 480 |
| QY | 491 | CTACCGGCTCATCGGGGTGGTGCTCTCTCCCTCATTTGACCCTGGTGGTTTGGCTCATGCTCTC | 550 |
| Db | 481 | CTACCGGCTCATCGGGGTGGTGCTCTCTCTCTCATTTGACCCTGGTGGTTTGGCTCATGCTCTC | 540 |
| QY | 551 | TTCCAGTTCTGGCAGGGCCACACAGGGATCAGGTACAAAGAGCAGAGGGAGAGCTGTGCC | 610 |
| Db | 541 | TTCCAGTTCTGGCAGGGCCACACAGGGATCAGGTACAAAGAGCAGAGGGAGAGCTGTGCC | 600 |
| QY | 611 | AAGGACGCTGTTCGCTGTGACGGGGGTGGTGGACCTGGCAAGCTGAAGAGTGAAGACCTGGCC | 670 |
| Db | 601 | AAGGACGCTGTTCGCTGTGACGGGGGTGGTGGACCTGGCAAGCTGAAGAGTGAAGACCTGGCC | 660 |
| QY | 671 | TGCGTGAAGTTGACTGCGGAGCAAGTCTCTGCTTAAATCTACTCTGGGTCTCCCATCAG | 730 |
| Db | 661 | TGCGTGAAGTTGACTGCGGAGCAAGTCTCTGCTTAAATCTACTCTGGGTCTCCCATCAG | 720 |
| QY | 731 | TGGCTTCCCATCTGTAGCAGCACTGGAATGACTCTTACTGAGAGAAACCTGCGACAG | 790 |
| Db | 721 | TGGCTTCCCATCTGTAGCAGCACTGGAATGACTCTTACTGAGAGAAACCTGCGACAG | 780 |
| QY | 791 | CTGGGTTTCGAGAGTGTCTCACCGGACAAACGAGGTGGCCACAGGATTTTGGCCAAAGC | 850 |
| Db | 781 | CTGGGTTTCGAGAGTGTCTCACCGGACAAACGAGGTGGCCACAGGATTTTGGCCAAAGC | 840 |
| QY | 851 | TTCTCAATCTTGAGATACAACTCCACCATCCGAGGAAAGCCTCCACAGGTCTGAATGCCCT | 910 |
| Db | 841 | TTCTCAATCTTGAGATACAACTCCACCATCCGAGGAAAGCCTCCACAGGTCTGAATGCCCT | 900 |
| QY | 911 | TCCGAGCGGATATCTCCCTCAAGTGTCCCACTGGGACTGAGGGCCATGACCGGGCGG | 970 |
| Db | 901 | TCCGAGCGGATATCTCCCTCAAGTGTCCCACTGGGACTGAGGGCCATGACCGGGCGG | 960 |
| QY | 971 | ATCGTGGGAGGGGGCGGTGGCTTCGGAATGCAAGTGGCTTTGGCAATGATCTGTCACTTC | 1030 |
| Db | 961 | ATCGTGGGAGGGGGCGGTGGCTTCGGAATGCAAGTGGCTTTGGCAATGATCTGTCACTTC | 1020 |
| QY | 1031 | GGCACCAACCCACATCTGTGAGAGGCACGCTCATTTGACGCGCCAGTGGGTGCTCACTGGCGCC | 1090 |
| Db | 1021 | GGCACCAACCCACATCTGTGAGAGGCACGCTCATTTGACGCGCCAGTGGGTGCTCACTGGCGCC | 1080 |
| QY | 1091 | CAGTGCCTTCTTCGTGACCCCGGGAGAGAGGTCTCTGAGGGCTGTGAAAGTGTGACGCGGACCC | 1150 |
| Db | 1081 | CAGTGCCTTCTTCGTGACCCCGGGAGAGAGGTCTCTGAGGGCTGTGAAAGTGTGACGCGGACCC | 1140 |
| QY | 1151 | AGCAACTGTGACCAAGTTGCTCTGAGGACGCTTCATTCGCGATCATATCAACGCAAT | 1210 |
| Db | 1141 | AGCAACTGTGACCAAGTTGCTCTGAGGACGCTTCATTCGCGATCATATCAACGCAAT | 1200 |
| QY | 1211 | TACACCGATGAGGAGGACGACATATGACATGGCCCATGCGGGCTGTCCAAACCCCTGAGC | 1270 |
| Db | 1201 | TACACCGATGAGGAGGACGACATATGACATGGCCCATGCGGGCTGTCCAAACCCCTGAGC | 1260 |
| QY | 1271 | CTGTCCGCTCATCTCACCCCTGTCTTCCCTCCCATGTGACAGACCTTTAGCTCAGT | 1330 |
| Db | 1261 | CTGTCCGCTCATCTCACCCCTGTCTTCCCTCCCATGTGACAGACCTTTAGCTCAGT | 1320 |
| QY | 1331 | GAGACCTGCTGGATTCACAGGCTTTTGGCAAGACGAGGAGACAGATGACAAAGACATCCCCC | 1390 |
| Db | 1321 | GAGACCTGCTGGATTCACAGGCTTTTGGCAAGACGAGGAGACAGATGACAAAGACATCCCCC | 1380 |
| QY | 1391 | TTCTCCCGGGAGGTGAGAGGTCAATCTCATGACCTTCAAGAAATGCAATGACTGTGGTTC | 1450 |
| Db | 1381 | TTCTCCCGGGAGGTGAGAGGTCAATCTCATGACCTTCAAGAAATGCAATGACTGTGGTTC | 1440 |
| QY | 1451 | TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGCCTTCTGTGGGGGCGAGAGAC | 1510 |
| Db | 1441 | TATGACAGTTACCTTACCCCAAGATGATGTGTGTGCTGGGAGCCTTCTGTGGGGGCGAGAGAC | 1500 |
| QY | 1511 | TCCTGCGAGGAGACAGCGGGGGGCTCTTGTGTGTGACGAAACACCGCTGTACTGTG | 1570 |

| | | | |
|----------|----------|---|------|
| Db | 1501 | TCCTGCCAGGAGACAGCGGGGGGCTTCTTGTGTGAGCAAGAAACACCGCTGTGTAACCTG | 1566 |
| Oy | 1571 | GCAGGTGTACACGCTGGGGGACACAGGCTGTGAGCAAGAAACAAACCTGTGTGTACACC | 1633 |
| Db | 1561 | GCAGGTGTCAACGAGTGGGGACACAGGCTGTGAGCAAGAAACAAACCTGTGTGTACACC | 1622 |
| Oy | 1611 | AAAGTGACAGAAAGTTCTTCCCTGTGATTTACAGCAAGATGAGAGGAGAGTGGATTCAG | 1691 |
| Db | 1621 | AAAGTGACAGAAAGTTCTTCCCTGTGATTTACAGCAAGATGAGAGGAGAGTGGATTCAT | 1680 |
| Oy | 1691 | AAATCCTAA 1699 | |
| Db | 1681 | AAATCCTAA 1689 | |
| RESULT 6 | | | |
| ADJ46916 | | | |
| ID | ADJ46916 | standard; CDNA; 1689 BP. | |
| XX | AC | ADJ46916; | |
| XX | DT | 06-MAY-2004 (first entry) | |
| XX | DE | Human transmembrane serine protease (MTSP)-related CDNA #2. | |
| KW | XX | Human; transmembrane serine protease; WTSP; gene; 88; | |
| KW | XX | cell surface protease; plasmin; prostate specific antigen; PSA; | |
| KW | XX | proliferative disease; cell-surface protease-associated disease; | |
| KW | XX | autoimmune disease; inflammatory disease; infectious disease; | |
| KW | XX | endocrine disease; cancer; ocular disorder; cardiovascular disorder; | |
| KW | XX | chronic inflammatory disease; wound; circulatory disorder; | |
| KW | XX | dermatological disorder; restenosis; rheumatoid arthritis; psoriasis; | |
| KW | XX | diabetic retinopathy; laser surgery; scarring; | |
| KW | XX | glaucoma filtering surgery scarring; macular degeneration; | |
| KW | XX | CRST syndrome; bacterial infection; viral disease; solid neoplasm; | |
| KW | XX | vascular tumour; lung; colon; prostate; melanoma; Kaposi's sarcoma. | |
| OS | XX | Homo sapiens. | |
| XX | PN | US2004001801-A1. | |
| XX | XX | 01-JAN-2004. | |
| XX | PD | 23-MAY-2002; 2002US-00156214. | |
| XX | PF | 23-MAY-2002; 2002US-00156214. | |
| XX | PR | 23-MAY-2002; 2002US-00156214. | |
| XX | PA | (CORV-) CORVAS INT INC. | |
| PI | XX | Madison EJ, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV; | |
| DR | XX | WPI; 2004-190126/18. | |
| PT | XX | P-PSDB; ADJ46917. | |
| XX | XX | Conjugate useful for treating e.g. cancer, cell-surface protease- | |
| XX | XX | associated diseases, comprising a peptidic substrate or nucleic acid | |
| XX | XX | substrate linked to a therapeutic agent through a linker. | |
| PS | XX | Claim 9; SEQ ID NO 23; 361pp; English. | |
| XX | XX | The invention relates to a conjugate comprising a therapeutic agent and a | |
| CC | XX | peptidic substrate or nucleic acid substrate linked to the agent | |
| CC | XX | optionally through a linker or peptidic linker, where the peptidic | |
| CC | XX | substrate is proteolytically cleaved by a cell surface protease or a | |
| CC | XX | soluble, released or shed form conjugate to liberate the agent and the | |
| CC | XX | conjugate is not substantially cleaved by plasmin or prostate specific | |
| CC | XX | antigen (PSA). The conjugate is useful for treating a disease, which | |
| CC | XX | involves administering a conjugate to a subject, where the disease is | |
| CC | XX | proliferative disease or a cell-surface protease-associated | |
| CC | XX | disease. The diseases include autoimmune diseases, inflammatory diseases, | |
| CC | XX | infectious diseases and endocrine diseases. The conjugate is useful for | |
| CC | XX | treating a cell-surface protease-associated disease, which involves | |

CC administering a conjugate comprising an agent and a peptidic substrate to
CC a subject exhibiting symptoms of a cell surface protease-associated
CC disorder, where the disease is selected from cancer, ocular disorders,
CC cardiovascular disorders, chronic inflammatory diseases, wounds,
CC circulatory disorders, dermatological disorders, restenosis, rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, scarring from laser
CC surgery, scarring from glaucoma filtering surgery, macular degeneration,
CC CREST syndrome, bacterial infections, viral diseases, solid neoplasms and
CC vascular tumours such as lung cancer, colon cancer, prostate cancer,
CC melanoma and Kaposi's sarcoma. This sequence represents cDNA encoding a
CC transmembrane serine protease (MTSP)-related polypeptide of the
CC invention.

SQ Sequence 1689 BP; 367 A; 529 C; 470 G; 323 T; 0 U; 0 Other;

| | | | | |
|-----------------------|--------------|---------------|---------------|--------------|
| Query Match | 96.5% | Score 1687.4; | DB 12; | Length 1689; |
| Best Local Similarity | 99.9% | Pred. No. 0; | | |
| Matches 1688; | Conservative | 0; | Mismatches 1; | Indels 0; |
| | | | Gaps | 0; |

| | | | | | | | | | | | | | | |
|----|-----|------------|-----------|---------|----------|----------|---------|--------|----------|---------|--------|-----------|-------|-----|
| QY | 1 | ATGAGAGGGA | CAGCC | CAGGAA | TGATCT | CAGCA | MAA | CACTT | CAGCTG | AGCATCT | 70 | | | |
| Db | 1 | ATGAGAGGGA | CAGCC | CAGGAA | TGATCT | CAGCA | MAA | CACTT | CAGCTG | AGCATCT | 60 | | | |
| QY | 71 | CCAGCC | CAGGCAT | CTTCC | CAGCTG | GGACA | CACTTCC | CAGGC | CGGGCAT | CTTCC | AGCC | CAGGCATCT | 130 | |
| Db | 61 | CCAGCC | CAGGCAT | CTTCC | CAGCTG | GGACA | CACTTCC | CAGGC | CGGGCAT | CTTCC | AGCC | CAGGCATCT | 120 | |
| QY | 131 | CCAGCC | CAGGCAT | CTTCC | CAGCTG | GGACA | CACTTCC | CAGGC | CGGGCAT | CTTCC | AGCC | CAGGCATCT | 190 | |
| Db | 121 | CCAGCC | CAGGCAT | CTTCC | CAGCTG | GGACA | CACTTCC | CAGGC | CGGGCAT | CTTCC | AGCC | CAGGCATCT | 180 | |
| QY | 191 | CCAGCTG | GTAACT | CTCC | CAGGCG | GGGATCT | CC | CAGGC | CGGGCAT | CTTCC | AGCC | CAGGCATCT | 250 | |
| Db | 181 | CCAGCTG | GTAACT | CTCC | CAGGCG | GGGATCT | CC | CAGGC | CGGGCAT | CTTCC | AGCC | CAGGCATCT | 240 | |
| QY | 251 | CCAGCC | CGGGCAT | CTTCC | CGGCT | CTGGCAT | CACTTTC | CAGGCT | CTCAT | CCGCG | CAGGCT | CACTCA | 310 | |
| Db | 241 | CCAGCC | CGGGCAT | CTTCC | CGGCT | CTGGCAT | CACTTTC | CAGGCT | CTCAT | CCGCG | CAGGCT | CACTCA | 300 | |
| QY | 311 | TCCG | CAGGTC | CAGCT | CGGTGA | CAACTTCC | CCCA | CAGAG | GTGATCTT | GTAG | CAACA | 370 | | |
| Db | 301 | TCCG | CAGGTC | CAGCT | CGGTGA | CAACTTCC | CCCA | CAGAG | GTGATCTT | GTAG | CAACA | 360 | | |
| QY | 371 | CCAGTGG | GGGCTGT | ATCCAT | CCGATCAT | TCTCTG | CCAG | GTGAG | CA | GCAC | CAGCA | CCA | GGGCC | 430 |
| Db | 361 | CCAGTGG | GGGCTGT | ATCCAT | CCGATCAT | TCTCTG | CCAG | GTGAG | CA | GCAC | CAGCA | CCA | GGGCC | 420 |
| QY | 431 | ACCA | GGGAGAGCC | CAGGTAC | AGAGCT | CGCCAG | TTC | CACTG | CGGG | AGAGG | CCAA | ACAG | 490 | |
| Db | 421 | ACCA | GGGAGAGCC | CAGGTAC | AGAGCT | CGCCAG | TTC | CACTG | CGGG | AGAGG | CCAA | ACAG | 480 | |
| QY | 491 | CTAC | CGCTCAT | TGGGG | ACGGT | CGCTCC | CTCAT | TGCG | CTGG | GTG | TTGG | CTCAT | CTC | 550 |
| Db | 481 | CTAC | CGCTCAT | TGGGG | ACGGT | CGCTCC | CTCAT | TGCG | CTGG | GTG | TTGG | CTCAT | CTC | 540 |
| QY | 551 | TTCC | AGTTCT | TGG | CAGGG | CCACA | CAGGGAT | CA | AGGTACA | AGAG | CAGAGG | AGAGCT | GTCC | 610 |
| Db | 541 | TTCC | AGTTCT | TGG | CAGGG | CCACA | CAGGGAT | CA | AGGTACA | AGAG | CAGAGG | AGAGCT | GTCC | 600 |
| QY | 611 | AAGCA | CGCTGTT | CGCTGT | GA | CGGGGT | GTGTGA | CTGCA | AGCTG | AAAG | ATGAT | CGAGCT | GGGC | 670 |
| Db | 601 | AAGCA | CGCTGTT | CGCTGT | GA | CGGGGT | GTGTGA | CTGCA | AGCTG | AAAG | ATGAT | CGAGCT | GGGC | 660 |
| QY | 671 | TGGGT | AGGTTTGA | CTGGGA | CAAGTCTG | CTAAAT | CTACT | CTGGGT | CTCTCC | ATCAG | 730 | | | |
| Db | 661 | TGGGT | AGGTTTGA | CTGGGA | CAAGTCTG | CTAAAT | CTACT | CTGGGT | CTCTCC | ATCAG | 720 | | | |
| QY | 731 | TGGCTT | CCATCTGT | AGAGCA | ACTG | GAATG | ACTCTT | CACTG | AGAG | AAACCT | GTCC | AGCAG | 790 | |
| Db | 721 | TGGCTT | CCATCTGT | AGAGCA | ACTG | GAATG | ACTCTT | CACTG | AGAG | AAACCT | GTCC | AGCAG | 780 | |
| QY | 791 | CTGG | GTTCG | AGAT | CTCA | CCGGACA | CCAGGTT | GCC | CAGG | GAATTT | GTCC | CAAC | 850 | |

| | | | |
|----|------|---|------|
| Db | 781 | CTGGGTTTCGAGATGCTCAACCCGACCAACGAGGTTGCCACAGGATTTTGGCAACAGC | 840 |
| Qy | 851 | TTCTCAATCTTTGAGATACAACTCCACCATCCAGAAAGCCTCCACAGGTCTGAATGCCCT | 910 |
| Db | 841 | TTCTCAATCTTTGAGATACAACTCCACCATCCAGAAAGCCTCCACAGGTCTGAATGCCCT | 900 |
| Qy | 911 | TCCACAGGGATATATCTCCCTCCAGTGTTCCTCAGGGGACTGAGGGGCTATGACCGGGGGG | 970 |
| Db | 901 | TCCACAGGGATATATCTCCCTCCAGTGTTCCTCAGGGGACTGAGGGGCTATGACCGGGGGG | 960 |
| Qy | 971 | ATCTGGGAGGGGGCGCTGGCCTCGATAGCAAGGCGCTTTGGCAAGTACTTGCATCTTC | 1030 |
| Db | 961 | ATCTGGGAGGGGGCGCTGGCCTCGGATAGCAAGTGGCTTTGGCAAGTACTTGCATCTTC | 1020 |
| Qy | 1031 | GGACCAACCCACATCTGTGAGAGCAGCAGCTATTTAGCGCCAGTGGGTGCTCACTGCCGC | 1090 |
| Db | 1021 | GGACCAACCCACATCTGTGAGAGCAGCAGCTATTTAGCGCCAGTGGGTGCTCACTGCCGC | 1080 |
| Qy | 1091 | CACGTCTTCTTCGTGACCCCGGAGAAAGTCTTGGAGGGCTTGGAAAGTGTACCGGGGCAC | 1150 |
| Db | 1081 | CACGTCTTCTTCGTGACCCCGGAGAAAGTCTTGGAGGGCTTGGAAAGTGTACCGGGGCAC | 1140 |
| Qy | 1151 | AGCAACCTGCACGAGTTGGCTTGAGGAGCCCTCAATTCGCGAGATCATCATCAACAGCAAT | 1210 |
| Db | 1141 | AGCAACCTGCACGAGTTGGCTTGAGGAGCCCTCAATTCGCGAGATCATCATCAACAGCAAT | 1200 |
| Qy | 1211 | TACACCGATGAGAGAGAGCACTATGATAGCGCCCTCATGCGGCTGTCCAAAGCCCTTAGC | 1270 |
| Db | 1201 | TACACCGATGAGAGAGAGCACTATGATAGCGCCCTCATGCGGCTGTCCAAAGCCCTTAGC | 1260 |
| Qy | 1271 | CTGTCCGCTCATATCCACCTGTGCTTGCCTCCCATGATGATGACAGACCTTAGCTCAAT | 1330 |
| Db | 1261 | CTGTCCGCTCATATCCACCTGTGCTTGCCTCCCATGATGATGACAGACCTTAGCTCAAT | 1320 |
| Qy | 1331 | GAGACCTGCTGGATCAACAGGCTTTGGCAAGCCAGGAGACAGATGACAAACATATCCCC | 1390 |
| Db | 1321 | GAGACCTGCTGGATCAACAGGCTTTGGCAAGCCAGGAGACAGATGACAAACATATCCCC | 1380 |
| Qy | 1391 | TTCTCCGGGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGACTACTTGTGTC | 1450 |
| Db | 1381 | TTCTCCGGGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGACTACTTGTGTC | 1440 |
| Qy | 1451 | TATGACAGTTACCTTTACCCCAAGATGATGTGTCTGGGACCTTGGTGGGGCAGAGAC | 1510 |
| Db | 1441 | TATGACAGTTACCTTTACCCCAAGATGATGTGTCTGGGACCTTGGTGGGGCAGAGAC | 1500 |
| Qy | 1511 | TCTCTCCAGGAGACACCGGGGGGCTCTTGTCTGTGACGAAACAACCGCTGGTACTGT | 1570 |
| Db | 1501 | TCTCTCCAGGAGACACCGGGGGGCTCTTGTCTGTGACGAAACAACCGCTGGTACTGT | 1560 |
| Qy | 1571 | GCAGGTGACACACTGTGGGGGACAGGCTGTGGCCAGAGAAACAACTGTGTTATACCC | 1630 |
| Db | 1561 | GCAGGTGACACACTGTGGGGGACAGGCTGTGGCCAGAGAAACAACTGTGTTATACCC | 1620 |
| Qy | 1631 | AAAGTGAACAGAAATTCTTCCCTGTGATTTACAGCAAGATGAGAGCGAGTGTCCATTAGA | 1690 |
| Db | 1621 | AAAGTGAACAGAAATTCTTCCCTGTGATTTACAGCAAGATGAGAGCGAGTGTCCATTAGA | 1680 |
| Qy | 1691 | AAATCTTAA 1699 | |
| Db | 1681 | AAATCTTAA 1689 | |

```

RESULT 7
ADY50145
XX      ADY50145 standard; DNA; 1689 BP
XX
AC      ADY50145;
XX
XX      19-MAY-2005 (first entry)
XX
XX      Human endothelias-2S DNA.
XX

```

KM Antiarthritic; Antirheumatic; Antipsoriatic; Antidiabetic;
KM Ophthalmological; Cardiovascular-Gen.; Vulnary; Antiinflammatory;
KM Vascotropic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
KM rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
KM dermatological disease; cancer; neoplasm; endothelase-2; de; gene.
XX Homo sapiens.
OS
XX WO2005019270-A2.
PN
PD 03-MAR-2005.
XX
XX 12-AUG-2004; 2004WO-US026148.
PF
XX 14-AUG-2003; 2003US-0495005P.
PR 14-NOV-2003; 2003US-0520164P.
XX
XX (DYAX-) DYAX CORP.
PA (BEND-) DENDREON CORP.
PI
PI Madison EL, Nixon A;
XX
XX MPI: 2005-202609/21.
DR P-PSDB; ADY50146.
XX
XX Novel protein capable of inhibiting endothelase-2, useful for treating
PT or preventing angiogenesis related disorder e.g. cancer.
XX
XX Disclosure; SEQ ID NO 93; 157pp; English.
PS
XX The invention relates to an isolated protein (I) which comprises a heavy
XX chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
XX immunoglobulin variable domain sequence, where the first and second
XX immunoglobulin variable domain sequences from an antigen binding site
XX that specifically binds to human endothelase-2. (I) is useful for
XX detecting an endothelase or endothelase activity in a sample, for
XX modulating an activity of an RT2-expressing cell, for modulating
XX proteolysis, for killing or inhibiting growth of a cell, for detecting
XX endothelase in a subject, for modulating endothelase activity in a
XX subject, for treating or preventing a disorder characterized by unwanted
XX angiogenesis in a subject. The disorder is chosen from rheumatoid
XX arthritis, psoriasis, diabetic retinopathies, ocular disorder such as
XX pterygial recurrence, scarring excimer laser surgery and glaucoma
XX filtering surgery, cardiovascular disorders, chronic inflammatory
XX disorders, wound repair, circulatory disorders, crest syndromes,
XX dermatological disorders and cancers. The present sequence represents the
XX human endothelase-2L DNA.
XX
XX Sequence 1689 BP; 367 A; 529 C; 470 G; 323 T; 0 U; 0 Other;
XX
XX Query Match 96.5%; Score 1687.4; DB 14; Length 1689;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 11 ATGAGAGAGGAGCAGCCAGGAATGATCTCCAGCAAGACACCTTCAGCTGAGCATCT 70
DB 1 ATGAGAGAGGAGCAGCCAGGAATGATCTCCAGCAAGACACCTTCAGCTGAGCATCT 60
QY 71 CCAGCCAGGATCTCCAGCTGGAGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 130
DB 61 CCAGCCAGGATCTCCAGCTGGAGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 120
QY 131 CCAGCCAGGATCTCCAGCTGGAGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 190
DB 121 CCAGCCAGGATCTCCAGCTGGAGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 180
QY 191 CCAGCTGTATACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGACATCT 250
DB 181 CCAGCTGTATACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGACATCT 240
QY 251 CCAGCCCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCA 310
DB 241 CCAGCCCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCA 300

QY 311 TCCGCAAGTCAAGCTGCTGGAGACACCTCCCAACAGAGATGATCTGTTAGACAACA 370
DB 301 TCCGCAAGTCAAGCTGCTGGAGACACCTCCCAACAGAGATGATCTGTTAGACAACA 360
QY 371 CCAAGTGGGGGCTGTATCCCATCCGATCATCTCTGCGCAGGTACAGCACAGCAAGGGCC 430
DB 361 CCAAGTGGGGGCTGTATCCCATCCGATCATCTCTGCGCAGGTACAGCACAGCAAGGGCC 420
QY 431 ACCAGGAGAGCCAGGTAAGAGCTGAGCTCCCAAGTTCACTGCGGGAGGGCCAGAAACAG 490
DB 421 ACCAGGAGAGCCAGGTAAGAGCTGAGCTCCCAAGTTCACTGCGGGAGGGCCAGAAACAG 480
QY 491 CTACCGCTCATTCGGGTCGCTGCTCTTCCTGATTTGCTGCTCATCTCTC 550
DB 481 CTACCGCTCATTCGGGTCGCTGCTCTTCCTGATTTGCTGCTCATCTCTC 540
QY 551 TTCCAGTTCTGGCAGGGCCACACAGGGATCAAGGTACAAAGACAGAGAGAGCTGTCCC 610
DB 541 TTCCAGTTCTGGCAGGGCCACACAGGGATCAAGGTACAAAGACAGAGAGCTGTCCC 600
QY 611 AAGCAGCTGTGCTGCTGAGCAGGGGTGAGACTGCAAGCTGAAGAGTGAAGAGCTGGCC 670
DB 601 AAGCAGCTGTGCTGCTGAGCAGGGGTGAGACTGCAAGCTGAAGAGTGAAGAGCTGGCC 660
QY 671 TGCCTGAGGTTTGAAGTGGGCAAGTCTGCTTAAATCTACTGAGTCTCCCATCAG 730
DB 661 TGCCTGAGGTTTGAAGTGGGCAAGTCTGCTTAAATCTACTGAGTCTCCCATCAG 720
QY 731 TGGCTTCCCATCTGTACAGACCACTGGAATGATCTCTACTCAGAGAAAGACTGGCACAG 790
DB 721 TGGCTTCCCATCTGTACAGACCACTGGAATGATCTCTACTCAGAGAAAGACTGGCACAG 780
QY 791 CTGGGTTTCAGAGATGCTCACCGGACCAACGAGGTTGCCACAGGATTTTGCACAGC 850
DB 781 CTGGGTTTCAGAGATGCTCACCGGACCAACGAGGTTGCCACAGGATTTTGCACAGC 840
QY 851 TTCTCAATCTTGAATCAACTCCACCATCCAGAAAGCTTCCACAGTCTGAATGCCCT 910
DB 841 TTCTCAATCTTGAATCAACTCCACCATCCAGAAAGCTTCCACAGTCTGAATGCCCT 900
QY 911 TCCAGAGGATATATCTCCCTCCAGTGTCCCACTGCGGACCTGAGAGGACATGACCGGGCCG 970
DB 901 TCCAGAGGATATATCTCCCTCCAGTGTCCCACTGCGGACCTGAGAGGACATGACCGGGCCG 960
QY 971 ATGCTGGAGGGGGGCTGAGCTCCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGACTTC 1030
DB 961 ATGCTGGAGGGGGGCTGAGCTCCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGACTTC 1020
QY 1031 GGCACCAACCAATCTGTGAGAGGACGCTGATGAGCGCCAGAGGAGTCTGACTGCGCC 1090
DB 1021 GGCACCAACCAATCTGTGAGAGGACGCTGATGAGCGCCAGAGGAGTCTGACTGCGCC 1080
QY 1091 CACTGCTTTCTGAGACCCCGGAGAAAGTCTGAGAGGGCTGAGAGGTGTACGGGGGACC 1150
DB 1081 CACTGCTTTCTGAGACCCCGGAGAAAGTCTGAGAGGGCTGAGAGGTGTACGGGGGACC 1140
QY 1151 AGCAACCTGACAGCTGTGAGAGGAGCTTCCATTCGAGATCATCATCAGCAAT 1210
DB 1141 AGCAACCTGACAGCTGTGAGAGGAGCTTCCATTCGAGATCATCATCAGCAAT 1200
QY 1211 TACACGATGAGAGAGAGACTATGATGAGCTGAGCTGAGAGCTGTCAGAGCCCTGACC 1270
DB 1201 TACACGATGAGAGAGAGACTATGATGAGCTGAGCTGAGAGCTGTCAGAGCCCTGACC 1260
QY 1271 CTGTCGCTCATCATCCACCTGCTGCTGCTCCCATGATGAGAGAGACTTTAGCTCAAT 1330
DB 1261 CTGTCGCTCATCATCCACCTGCTGCTGCTCCCATGATGAGAGAGACTTTAGCTCAAT 1320
QY 1331 GAGACCTGCTGATCAAGGCTTTGGCAAGACAGGAGACAGATGACAAGACATCTCCCC 1390
DB 1321 GAGACCTGCTGATCAAGGCTTTGGCAAGACAGGAGACAGATGACAAGACATCTCCCC 1380

QY 1391 TTCCTCCGGAGGAGTGCATCATCTCATGACTTCAAGAAATGCAATGACTTGTGTC 1450
 XX
 Db 1381 TTCTTCGGAGGAGTGCATCATCTCATGACTTCAAGAAATGCAATGACTTGTGTC 1440
 QY 1451 TATGACAGTTACCTTACCTCCCAAGAGATGATGTGCTGGGAGACTTGTGGGGGAGAGAC 1510
 Db 1441 TATGACAGTTACCTTACCTCCCAAGAGATGATGTGCTGGGAGACTTGTGGGGGAGAGAC 1500
 QY 1511 TCTCTCCAGGAGACAGCGGGGGGCTTGTGTCTGTGTGACGAACAACCGTGTGACTGTG 1570
 Db 1501 TCTCTCCAGGAGACAGCGGGGGGCTTGTGTCTGTGTGACGAACAACCGTGTGACTGTG 1560
 QY 1571 GCAGGTTCACTACCTGAGGAGGAGGCTGTGTGTGACGAACAACCGTGTGACTGTG 1630
 Db 1561 GCAGGTTCACTACCTGAGGAGGAGGCTGTGTGTGACGAACAACCGTGTGACTGTG 1620
 QY 1631 AAAGTGAAGAGAGTTCTTCTCTGATTTTACAGCAAGATGAGAGCGAGTGTGATTCAGA 1690
 Db 1621 AAAGTGAAGAGAGTTCTTCTCTGATTTTACAGCAAGATGAGAGCGAGTGTGATTCATA 1680
 QY 1691 AAATCTTAA 1699
 Db 1681 AAATCTTAA 1689

RESULT 8
 AD275551
 ID AD275551 standard; cDNA; 2380 BP.

XX AD275551;

DT 14-JUL-2005 (first entry)

XX Human mosaic serine protease (MSP) cDNA.

XX
 KW cardiant; endocrine-gen.; metabolic; gastrointestinal; hemostatic;
 KW respiratory-gen.; nootropic; neuroprotective; uropathic; gynecological;
 KW antiinflammatory; cytostatic; antianemic; diagnostic; pharmaceutical;
 KW cardiovascular disease; cardiovascular-gen.; endocrine-gen.; metabolic;
 KW metabolic disorder; hematological disease; respiratory disease;
 KW respiratory-gen.; neurological disease; andrology; endocrine disease;
 KW gastrointestinal disease; gynecology and obstetrics; inflammation;
 KW cancer; cytostatic; neoplasm; mosaic serine protease; MSP; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 88..1833

XX /product /transl_except= (pos:862..864, aa:Arg)

XX /transl_except= (pos:979..981, aa:His)

XX W02005040401-A2.

XX 06-MAY-2005.

XX 02-OCT-2004; 2004WO-EP011015.

XX 17-OCT-2003; 2003EP-00023809.

XX (FARB) BAYER HEALTHCARE AG.

XX Golz S, Bruegemeier U, Geerts A;

XX WPI: 2005-355849/36.

XX P-PSDB; AD275552.

XX Screening for therapeutic agents for treating cardiovascular, metabolic,
 PT hematological, neurological, urological or respiratory diseases, or
 PT inflammation by contacting a test compound with a mosaic serine protease
 PT polypeptide.

PS Disclosure; SEQ ID NO 1; 98pp; English.

XX The invention describes a method of screening for therapeutic agents for
 CC treating cardiovascular, endocrinological, metabolic, gastrointestinal,
 CC hematological, respiratory, neurological, urological and reproduction
 CC disorders, inflammation, or cancer in a mammal comprising contacting a
 CC test compound with a mosaic serine protease (MSP) polypeptide. The method
 CC comprises: contacting a test compound with a MSP polypeptide; and
 CC detecting binding of the test compound to the MSP polypeptide. Also
 CC described are: a method of diagnosing any of the diseases cited above in
 CC a mammal comprising determining the amount of a MSP polynucleotide in a
 CC sample taken from the mammal; and determining the amount of MSP
 CC polynucleotide in healthy and/or diseased mammals; a pharmaceutical
 CC composition for the treatment of any of the diseases cited above,
 CC comprising a therapeutic agent which binds to, or regulates the activity
 CC of, a MSP polypeptide; or a MSP polynucleotide or polypeptide; a method
 CC for preparing the pharmaceutical composition comprising identifying a
 CC regulator of MSP; determining whether the regulator ameliorates the
 CC symptoms the disease in a mammal; and combining of the regulator with an
 CC acceptable pharmaceutical carrier; and a regulator of MSP for the
 CC regulation of MSP activity in a mammal having any of the diseases cited
 CC above. The method is useful for screening for therapeutic agents for
 CC treating cardiovascular, endocrinological, metabolic, gastrointestinal,
 CC hematological, respiratory, neurological, urological and reproduction
 CC disorders, inflammation, or cancer in a mammal. The compositions are
 CC useful for treating such diseases. This sequence encodes human mosaic
 CC serine protease (MSP).

XX Sequence 2380 BP; 540 A; 692 C; 627 G; 521 T; 0 U; 0 Other;

XX Query Match 95.6%; Score 1670.4; DB 14; Length 2380;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCAGAGACCATGAGAGGAGACCCAGCGGAATGATCTTCAGCAAGAACCTTCCAGC 60
 Db 78 CTCAGAGACCATGAGAGGAGAGACCCAGCGGAATGATCTTCAGCAAGAACCTTCCAGC 137
 QY 61 TGGAGCATCTTCAGCCCGGAGATCTCCAGCTGGAGACCTTCAGGCCGGGATCTCCAGC 120
 Db 138 TGGAGCATCTTCAGCCCGGAGATCTCCAGCTGGAGACCTTCAGGCCGGGATCTCCAGC 197
 QY 121 CCAAGCATCTTCAGCCCGGAGATCTCCAGCTGGAGACCTTCAGGCCGGGATCTCCAGC 180
 Db 198 CCAAGCATCTTCAGCCCGGAGATCTCCAGCTGGAGACCTTCAGGCCGGGATCTCCAGC 257
 QY 181 CCAAGCATCTTCAGCTGGTACCTTCAGGCCCGGAGATCTTCAGGCCGGGATCTCCAGC 240
 Db 258 CCAAGCATCTTCAGCTGGTACCTTCAGGCCCGGAGATCTTCAGGCCGGGATCTCCAGC 317
 QY 241 CCAAGCATCTTCAGCCCGGAGATCTCCAGCTGGAGACCTTCAGGCCGGGATCTCCAGC 300
 Db 318 CCAAGCATCTTCAGCCCGGAGATCTCCAGCTGGAGACCTTCAGGCCGGGATCTCCAGC 377
 QY 301 CAGGTATCATTCGCGCAGGTGAGCTGGTGAACAACCTCCCAACGAGATGATCTTGT 360
 Db 378 CAGGTATCATTCGCGCAGGTGAGCTGGTGAACAACCTCCCAACGAGATGATCTTGT 437
 QY 361 TAGAGCAACACCAAGTGGGGCTGTACCATTCGATCTTCCTGCGAGGTCCAGCAGC 420
 Db 438 TAGAGCAACACCAAGTGGGGCTGTACCATTCGATCTTCCTGCGAGGTCCAGCAGC 497
 QY 421 AACCAAGGCAACCAAGGAGAGCCAGGTAGAGCTTCGCCAAGTTCACTTGGCGGAGGG 480
 Db 498 AACCAAGGCAACCAAGGAGAGCCAGGTAGAGCTTCGCCAAGTTCACTTGGCGGAGGG 557
 QY 481 CCAAGAGCAGTACCGGTATCGGGTGGCTCTCTCTCATTTGCGTGGTGGTTGGCT 540
 Db 558 CCAAGAGCAGTACCGGTATCGGGTGGCTCTCTCTCATTTGCGTGGTGGTTGGCT 617
 QY 541 CATCATCTCTTCCAGTTCTGGAGGGCCACAGAGGATCAGGTACAGAGACAGAGGGA 600
 Db 618 CATCATCTCTTCCAGTTCTGGAGGGCCACAGAGGATCAGGTACAGAGACAGAGGGA 677

QY 601 GAGCTGTCCCAAGCAGCTGTTCCTGTGACGCGGTGTGTGACCTGCAAGCTGAAGATGA 660
 DB 678 GAGCTGTCCCAAGCAGCTGTTCCTGTGACGCGGTGTGTGACCTGCAAGCTGAAGATGA 737
 QY 661 CGAGCTGGGCTGCGTGAAGTTTGACTGGGACAACTCTCTGTTAAATCTACTCTGGGTC 720
 DB 738 CGAGCTGGGCTGCGTGAAGTTTGACTGGGACAACTCTCTGTTAAATCTACTCTGGGTC 797
 QY 721 CTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAATGACTCTACTCAGAGAAAGAC 780
 DB 798 CTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAATGACTCTACTCAGAGAAAGAC 857
 QY 781 CTGCGCAGCAGCTGGGTTTCCGAGTGTCTCAGCGGACAACTGAGTTGCCCAAGGATTT 840
 DB 858 CTGCGCAGCAGCTGGGTTTCCGAGTGTCTCAGCGGACAACTGAGTTGCCCAAGGATTT 917
 QY 841 TGGCAACAGCTTCTCAATCTTGAGATACAACTCCACCACTCCAGAAAGCTCCACAGTTC 900
 DB 918 TGGCAACAGCTTCTCAATCTTGAGATACAACTCCACCACTCCAGAAAGCTCCACAGTTC 977
 QY 901 TGAATGCCCTTCCGAGCGGTATATCTCCCTCAGTGTTCCTCAGCGGACTGAGGCGCAT 960
 DB 978 TGAATGCCCTTCCGAGCGGTATATCTCCCTCAGTGTTCCTCAGCGGACTGAGGCGCAT 1037
 QY 961 GACCGGCGGATCGTGGAGGGGCGCTGGCTCGGATGACAAAGTGGCTTGGCAAGTGAG 1020
 DB 1038 GACCGGCGGATCGTGGAGGGGCGCTGGCTCGGATGACAAAGTGGCTTGGCAAGTGAG 1097
 QY 1021 TCGCACTTGGGGACCAACCACTGTGTGAGAGGACCTCACTTACCGCCAGTGGGTGCT 1080
 DB 1098 TCGCACTTGGGGACCAACCACTGTGTGAGAGGACCTCACTTACCGCCAGTGGGTGCT 1157
 QY 1081 CACTGCGCGCCACTGTCTTCTGTGACCCGCGGAGAAAGTCTCTGAGAGGCTTGAAGTGTGA 1140
 DB 1158 CACTGCGCGCCACTGTCTTCTGTGACCCGCGGAGAAAGTCTCTGAGAGGCTTGAAGTGTGA 1217
 QY 1141 CGCGGCGACCAAGCACTGACCAAGTGTCTGAGGCAAGCTTCACTTGGCCAGATCATCAT 1200
 DB 1218 CGCGGCGACCAAGCACTGACCAAGTGTCTGAGGCAAGCTTCACTTGGCCAGATCATCAT 1277
 QY 1201 CAACAGCAATTAACCCGATGAGAGGAGCACTATGACATGCGCTCATGCGGCTGTCCAA 1260
 DB 1278 CAACAGCAATTAACCCGATGAGAGGAGCACTATGACATGCGCTCATGCGGCTGTCCAA 1337
 QY 1261 GCGCCCTGACCTGTCCGCTCATCACTCACTGCTTGGCTCCCAATGATGACAGACCTT 1320
 DB 1338 GCGCCCTGACCTGTCCGCTCATCACTCACTGCTTGGCTCCCAATGATGACAGACCTT 1397
 QY 1321 TAGCTCAATGAGACCTGTGTGATCAAGGCTTTGGCAAGCAAGGAGACAGATGACAA 1380
 DB 1398 TAGCTCAATGAGACCTGTGTGATCAAGGCTTTGGCAAGCAAGGAGACAGATGACAA 1457
 QY 1381 GACATCCCTCTCTCCGCGGAGGTGAGGTCAATCTCATGCACTTCAAGAAATCCAAATGA 1440
 DB 1458 GACATCCCTCTCTCCGCGGAGGTGAGGTCAATCTCATGCACTTCAAGAAATCCAAATGA 1517
 QY 1441 CTACTTGTGTATGACAGTTACTTATCCCAAGATGATGTGTGTGGGAGCTTCTGTGG 1500
 DB 1518 CTACTTGTGTATGACAGTTACTTATCCCAAGATGATGTGTGTGGGAGCTTCTGTGG 1577
 QY 1501 GGGCAGAGACTCTGTGCAAGGAGACAGCGGGGGGCTTGTGTGTGTGAGCAAGCAACCG 1560
 DB 1578 GGGCAGAGACTCTGTGCAAGGAGACAGCGGGGGGCTTGTGTGTGTGAGCAAGCAACCG 1637
 QY 1561 CTGTACTCTGCAAGTGTCACTCACTGGGGGACAGGCTGTGGCCAGAGAAACAACTCTGG 1620
 DB 1638 CTGTACTCTGCAAGTGTCACTCACTGGGGGACAGGCTGTGGCCAGAGAAACAACTCTGG 1697
 QY 1621 TGTGTACACCAAGTGAAGAGTTTCTCTGTGATTTACAGAAAGATGAG 1672
 DB 1698 TGTGTACACCAAGTGAAGAGTTTCTCTGTGATTTACAGAAAGATGAG 1749

RESULT 9
 ADE31742
 ID ADE31742 standard; DNA; 2393 BP.
 XX
 AC ADE31742;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human 2914 gene #SEQ ID 99.
 XX
 KM Antiheterosclerotic; cardiact; vasotrophic; antiinflammatory;
 KM thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
 KM cardiovascular disorder; ischaemia; aortic bending;
 KM vascular heart disease; endocarditis; atrial fibrillation; heart failure;
 KM angina; cardiomyopathy; cardiac death; gene; de.
 XX
 OS Homo sapiens.
 XX
 PN W02003065984-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 29-JAN-2003; 2003MO-US002571.
 XX
 PR 01-FEB-2002; 2002US-0353224P.
 PR 15-MAR-2002; 2002US-0364529P.
 PR 19-APR-2002; 2002US-0373611P.
 PR 29-APR-2002; 2002US-0376287P.
 PR 12-JUN-2002; 2002US-0388080P.
 PR 24-JUN-2002; 2002US-0390971P.
 PR 03-JUL-2002; 2002US-0394110P.
 PR 10-JUL-2002; 2002US-0394797P.
 PR 21-AUG-2002; 2002US-0404904P.
 PR 23-AUG-2002; 2002US-0405450P.
 PR 04-SEP-2002; 2002US-0408070P.
 PR 06-NOV-2002; 2002US-0424300P.
 PR 05-DEC-2002; 2002US-0431042P.
 PR 05-DEC-2002; 2002US-0431079P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Logan TU, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
 PI Steglano N, Perodin J, Rodrigue-Way A;
 XX
 DR WPI; 2003-731468/69.
 DR P-PSDB; ADE31743.
 XX
 PT Identifying a compound capable of treating a cardiovascular disorder
 PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
 PT modulate the expression or activity of e.g. 1682, 6169 or 6193
 PT polypeptide or nucleic acid.
 XX
 PS Disclosure; SEQ ID NO 99; 328bp; English.
 XX
 CC The invention relates to a method for identifying a compound capable of
 CC treating a cardiovascular disorder. The present invention identifies the
 CC differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,
 CC 43126, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,
 CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,
 CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
 CC 2868, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
 CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or
 CC 6585 genes in cardiovascular disease states. The methods are useful in
 CC diagnosing, preventing and treating cardiovascular disorders, such as
 CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,
 CC resection, arterial inflammation, vascular wall remodeling, coronary
 CC microembolism, tachycardia, bradycardia, pressure overload, aortic
 CC bending, coronary artery ligation, vascular heart disease, valvular
 CC disease, including but not limited to, valvular degeneration caused by
 CC calcification, rheumatic heart disease, endocarditis, or complications of
 CC artificial valves, atrial fibrillation, long-QT syndrome, congestive
 CC heart failure, sinus node dysfunction, angina, heart failure,

CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,
 CC including but not limited to, pericardial effusion and pericarditis;
 CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
 CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary
 CC artery spasm, ischemic disease, arrhythmia, sudden cardiac death and
 CC cardiovascular developmental disorders. The methods may also be used for
 CC identifying compounds that modulate cardiovascular disorders. Sequences
 CC given in ADE31644-ADE31769 represent the genes and proteins that may be
 CC regulated by a compound of the invention.

XX Sequence 2393 BP; 553 A; 692 C; 627 G; 521 T; 0 U; 0 Other;

Query Match 95.6%; Score 1670.4; DB 10; Length 2393;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CTCAGAGACATGAGAGAGAGACAGCCAGGAAATGATCTCCAGACAGAACACCTTCAGC 60
DB CTCAGAGACATGAGAGAGAGAGACAGCCAGGAAATGATCTCCAGACAGAACACCTTCAGC 137
QY 61 TGGAGCATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 120
DB TGGAGCATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 197
QY 121 CCAGGCAATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 180
DB CCAGGCAATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 257
QY 181 CCAGGCAATCTCCAGCTGGTACACCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 240
DB CCAGGCAATCTCCAGCTGGTACACCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 317
QY 241 CCAGGCAATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 300
DB CCAGGCAATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 377
QY 301 CAGGTCATCATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 360
DB CAGGTCATCATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 437
QY 378 CAGGTCATCATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 437
DB CAGGTCATCATCTCCAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 497
QY 438 TAGAGCAACACCAATGGGGGCTGTACCATCCATCATCTCCAGCTGGGATCTCCAGC 420
DB TAGAGCAACACCAATGGGGGCTGTACCATCCATCATCTCCAGCTGGGATCTCCAGC 497
QY 438 TAGAGCAACACCAATGGGGGCTGTACCATCCATCATCTCCAGCTGGGATCTCCAGC 420
DB TAGAGCAACACCAATGGGGGCTGTACCATCCATCATCTCCAGCTGGGATCTCCAGC 497
QY 421 AACCAAGGAGCCAGGAGAGAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 480
DB AACCAAGGAGCCAGGAGAGAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 557
QY 498 AACCAAGGAGCCAGGAGAGAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 557
DB AACCAAGGAGCCAGGAGAGAGCCAGGAGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 557
QY 481 CCAGAGAGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGC 540
DB CCAGAGAGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGC 617
QY 541 CATCATCTCTCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGC 600
DB CATCATCTCTCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGC 677
QY 618 CATCATCTCTCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGC 600
DB CATCATCTCTCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCCAGC 677
QY 601 GAGCTGTCTCCAGAGAGAGCTGTCTGTGAGAGGGGTGTGTGAGAGAGAGAGAGAGAG 660
DB GAGCTGTCTCCAGAGAGAGCTGTCTGTGAGAGGGGTGTGTGAGAGAGAGAGAGAGAG 737
QY 678 GAGCTGTCTCCAGAGAGAGCTGTCTGTGAGAGGGGTGTGTGAGAGAGAGAGAGAGAG 737
DB GAGCTGTCTCCAGAGAGAGCTGTCTGTGAGAGGGGTGTGTGAGAGAGAGAGAGAGAG 797
QY 661 CGAGCTGGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB CGAGCTGGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 797
QY 721 CTCCCATCATGAGCTCTCCATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB CTCCCATCATGAGCTCTCCATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 857
QY 798 CTCCCATCATGAGCTCTCCATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 857
DB CTCCCATCATGAGCTCTCCATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 917
QY 858 CTGCGAGAGAGCTGGGTTTCGAGAGAGCTTCACCGGAGAGAGAGAGAGAGAGAGAGAG 917
DB CTGCGAGAGAGCTGGGTTTCGAGAGAGCTTCACCGGAGAGAGAGAGAGAGAGAGAGAG 900
QY 841 TGCCAGAGAGCTTTCATCTTGAATACCACTCCACATCCAGAGAGAGAGAGAGAGAG 900

```

```

DB 918 TGCCAGAGCTTTCATCTTGAATACCACTCCACATCCAGAGAGAGAGAGAGAGAG 977
QY 901 TGAATGCCCTTCCAGAGAGATATCTCTCCATGATTTCCCATCTGAGAGAGAGAGAT 960
DB TGAATGCCCTTCCAGAGAGATATCTCTCCATGATTTCCCATCTGAGAGAGAGAGAT 1037
QY 961 GAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB GAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097
QY 1038 GAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097
DB GAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1021 TCTGCACTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB TCTGCACTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157
QY 1098 TCTGCACTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157
DB TCTGCACTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
QY 1081 CACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB CACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
QY 1158 CACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
DB CACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277
QY 1141 CGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB CGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277
QY 1201 CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1337
QY 1278 CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1337
DB CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1261 GCGGTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB GCGGTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397
QY 1338 GCGGTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397
DB GCGGTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1321 TAGCCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB TAGCCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
QY 1398 TAGCCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
DB TAGCCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1381 GACATCCCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB GACATCCCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1517
QY 1458 GACATCCCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1517
DB GACATCCCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1441 CTACTTGTCTATACAGATTACCTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB CTACTTGTCTATACAGATTACCTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1577
QY 1518 CTACTTGTCTATACAGATTACCTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1577
DB CTACTTGTCTATACAGATTACCTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
QY 1501 GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1637
QY 1578 GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1637
DB GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1561 CTGCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB CTGCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697
QY 1638 CTGCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697
DB CTGCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
QY 1621 TGTGTACACCAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1672
DB TGTGTACACCAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1749
QY 1698 TGTGTACACCAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1749
DB TGTGTACACCAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```

RESULT 10

AAD05797 ID AAD05797 standard; DNA; 2067 BP.

XX AAD05797;

XX 31-JUL-2001 (first entry)

XX Human transmembrane serine protease (Endothelinase 2-L) DNA.

DE Human, endothelinase 2-L; protease domain; cytosolic; vulnere; wound;
 KW neotrophic; peridontitis; dermatological disorder; gene therapy; scar;
 KW angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
 KW chronic inflammatory disease; ocular disorder; circulatory disorder;
 KW crest syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
 KW liver cirrhosis; osteoradionecrosis; systemic sclerosis; oesophagel;
 KW inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;

Db 1141 AGCAACCTGCACGATTGCTGAGGACGCTCCATGTCGCCGATCATCATCAACAGCAAT 1200
Qy 1211 TACACCGATGAGGAGGACGATATGACATGCGCCCTCATGCGGCTGTCCAGCCCTTGACC 1270
Db 1201 TACACCGATGAGGAGGACGATATGACATGCGCCCTCATGCGGCTGTCCAGCCCTTGACC 1260
Qy 1271 CTGACCGGTCATATCCACCCCTGCTGCTCCCAATGATGAGACAGACCTTTAGCCTCAAT 1330
Db 1261 CTGACCGGTCATATCCACCCCTGCTGCTCCCAATGATGAGACAGACCTTTAGCCTCAAT 1320
Qy 1331 GAGACCTGCTGATACAGAGCTTTGGCAGACAGAGAGACAGATGACAGACATCCCCC 1390
Db 1321 GAGACCTGCTGATACAGAGCTTTGGCAGACAGAGAGACAGATGACAGACATCCCCC 1380
Qy 1391 TTCTCTCCGGGAGGTCAGAGTCAATCTCATGCACTTCAAGAAATGCAATGATTAATTGCTC 1450
Db 1381 TTCTCTCCGGGAGGTCAGAGTCAATCTCATGCACTTCAAGAAATGCAATGATTAATTGCTC 1440
Qy 1451 TATGACAGTTACCTTACCCCAAGATGATGTCGTCGGGACCTTCGTCGGGAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGATGATGTCGTCGGGACCTTCGTCGGGAGAGAC 1500
Qy 1511 TCCTGCGACGAGGACAGCGGGGGCTCTGTCGTGTGAGCAGAAACAACCGCTGTACCTG 1570
Db 1501 TCCTGCGACGAGGACAGCGGGGGCTCTGTCGTGTGAGCAGAAACAACCGCTGTACCTG 1560
Qy 1571 GCAGGTGTACACAGCTGGGGGACAGGCTGTGTGGCCAGAGAAACAACCTGTGTGTACACC 1630
Db 1561 GCAGGTGTACACAGCTGGGGGACAGGCTGTGTGGCCAGAGAAACAACCTGTGTGTACACC 1620
Qy 1631 AAAATGACAGAAATTTCTTCCCTGATTTTACAGCAAGATGAGAGC 1675
Db 1621 AAAATGACAGAAATTTCTTCCCTGATTTTACAGCAAGATGAGAGC 1665

RESULT 11
AD110394
ID AD110394 standard; DNA; 2067 BP.
AC AD110394;
XX
DT 22-APR-2004 (first entry)
DE Human cell surface protease coding sequence #13.
XX
KW therapeutic agent; plasmin; protease specific antigen; PSA;
KW cell-surface protease-associated disease; cancer; ocular disease;
KW cardiovascular disease; chronic inflammatory disease; wound;
KW circulatory disorder; dermatological disorder; rheumatoid arthritis;
KW psoriasis; diabetic retinopathy; pterygium;
KW excimer laser surgery scarring; glaucoma filtering surgery scarring;
KW macular degeneration; crest syndrome; solid neoplasm; vascular tumour;
KW melanoma; Kaposi's sarcoma; human; cell surface protease; gene; ds.
XX
OS Homo sapiens.
XX
FN WO200295007-A2.
XX
PD 28-NOV-2002.
XX
PF 23-MAY-2002; 2002WO-US016819.
XX
PR 23-MAY-2001; 2001US-0293267P.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Madison EL, Semple JB, Vlasuk GP, Kemp SU, Komandla M, Siev DV;
XX
DR WPI; 2003-221280/21.
XX
DR P-PSDB; AD110395.
XX
PT Novel conjugate useful for treating cell-surface protease-associated
PT disease, comprises a therapeutic agent and a peptidic or nucleic acid

PT substrate linked to it optionally by a peptidic linker.
XX
PS Claim 9; SEQ ID NO 25; 581bp; English.
XX
CC The invention comprises a conjugate that consists of a therapeutic agent
CC and a peptide substrate (optionally linked via linker). The peptide
CC substrate is proteolytically cleaved by a cell surface protease pr a
CC soluble, released or shed form of it, to liberate the therapeutic agent,
CC the conjugate of the invention is not substantially cleaved by plasmin or
CC protease specific antigen (PSA). The conjugate of the invention is useful
CC for treating a cell-surface protease-associated disease such as: cancer;
CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
CC wounds, circulatory disorders, dermatological disorders, rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
CC scarring from excimer laser surgery, scarring from glaucoma filtering
CC surgery, macular degeneration, crest syndromes, solid neoplasms, vascular
CC tumours, melanoma and Kaposi's sarcoma. The present DNA sequence encodes
CC a human cell surface protease.
XX
SQ Sequence 2067 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;

Query Match 95.2%; Score 1663.4; DB 10; Length 2067;
Best Local Similarly 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 ATGAGAGGGGACAGCCACGGGAATGCATCTCCAGCAAGAACCTTCAGTGGAGCATCT 70
Db 1 ATGAGAGGGGACAGCCACGGGAATGCATCTCCAGCAAGAACCTTCAGTGGAGCATCT 60
Qy 71 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 130
Db 61 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 120
Qy 131 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 190
Db 121 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 180
Qy 191 CCAAGCTGTACACTCCAGGCGGGGCATCTCCAGGCGGGGCATCTCCAGCCAGGCATCT 250
Db 181 CCAAGCTGTACACTCCAGGCGGGGCATCTCCAGGCGGGGCATCTCCAGCCAGGCATCT 240
Qy 251 CCAAGCCCGGCATCTCCAGCTGGGACACCTCCAGCTTCATCTCCAGGCGGCATCTCA 310
Db 241 CCAAGCCCGGCATCTCCAGCTGGGACACCTCCAGCTTCATCTCCAGGCGGCATCTCA 300
Qy 311 TCAGCCAGGTACGCTTCGTGACAACTCCCAACCAAGGTATCTTTTAAAGCAACA 370
Db 301 TCAGCCAGGTACGCTTCGTGACAACTCCCAACCAAGGTATCTTTTAAAGCAACA 360
Qy 371 CCAAGTGGGGGTGTAACCATCCGATCATCTCTGCGAGGTGACAGCAGCAAGGAGCC 430
Db 361 CCAAGTGGGGGTGTAACCATCCGATCATCTCTGCGAGGTGACAGCAGCAAGGAGCC 420
Qy 431 ACCAGGAGAGCCACGAGTACGAGGCTCCCAAGTTCACTTGGCGGAGGGCCAGAAAGCAG 490
Db 421 ACCAGGAGAGCCACGAGTACGAGGCTCCCAAGTTCACTTGGCGGAGGGCCAGAAAGCAG 480
Qy 491 CTACCGCTCATCGGGTCGTGCTCTCTCTCATTTGCGCTGTGTTGCTCATATCTCTC 550
Db 481 CTACCGCTCATCGGGTCGTGCTCTCTCTCATTTGCGCTGTGTTGCTCATATCTCTC 540
Qy 551 TTCCAGTTCTGGGAGGAGCAACAGAGGATGACAGAGGAGAGAGAGCTGTCTCC 610
Db 541 TTCCAGTTCTGGGAGGAGCAACAGAGGATGACAGAGGAGAGAGAGCTGTCTCC 600
Qy 611 AAGCAGCTGTTCGCTGTGACGGGGTGTGAGCTGCAAGCTGAAGAGTGAAGAGTGGAGC 670
Db 601 AAGCAGCTGTTCGCTGTGACGGGGTGTGAGCTGCAAGCTGAAGAGTGAAGAGTGGAGC 660
Qy 671 TGGGTAGGTTTGAAGTGGAGCAAGTCTCTGTTAAATCTAATCTTGGGTCTCTCCATCAG 730
Db 661 TGGGTAGGTTTGAAGTGGAGCAAGTCTCTGTTAAATCTAATCTTGGGTCTCTCCATCAG 720

| | | | |
|----|------|--|------|
| Qy | 731 | TGGCTTCCCAATCTGTGAGAGCAACTGGAAATGACTCTCTACTCAGAGAAACAACCTGCCAGAG | 790 |
| Db | 721 | TGGCTTCCCAATCTGTGAGAGCAACTGGAAATGACTCTCTACTCAGAGAAACAACCTGCCAGAG | 780 |
| Qy | 791 | CTGGGTTTCGAGAGAGTCTCAACCGGCAACCGAGGTTGGCCACAGGGATTTTGGCCAAAGC | 850 |
| Db | 781 | CTGGGTTTCGAGAGAGTCTCAACCGGCAACCGAGGTTGGCCACAGGGATTTTGGCCAAAGC | 840 |
| Qy | 851 | TTCTCAATCTTGAGATACAACTCCACCATCGAGAAAGCCTTCCAGGTTGTAATGCCCT | 910 |
| Db | 841 | TTCTCAATCTTGAGATACAACTCCACCATCGAGAAAGCCTTCCAGGTTGTAATGCCCT | 900 |
| Qy | 911 | TCCCAAGCGGTATATCTCCCTCCAGTGTTCACATCGCGGATCTGAGGGCCATGACCGGGGG | 970 |
| Db | 901 | TCCCAAGCGGTATATCTCCCTCCAGTGTTCACATCGCGGATCTGAGGGCCATGACCGGGGG | 960 |
| Qy | 971 | ATCGTGGAGAGGAGCGCTGGCCTCGGATAGCAAGTGGCCTTGACCAATGAGTGTGCACTTC | 1038 |
| Db | 961 | ATCGTGGAGAGGAGCGCTGGCCTCGGATAGCAAGTGGCCTTGACCAATGAGTGTGCACTTC | 1020 |
| Qy | 1031 | GGACCAACCAATCTGTGAGAGCAAGCTCATTTGACGCCCAGTGGGTCTCATGCGGCC | 1090 |
| Db | 1021 | GGACCAACCAATCTGTGAGAGCAAGCTCATTTGACGCCCAGTGGGTCTCATGCGGCC | 1080 |
| Qy | 1091 | CACGTCTTCTTCGTGACCCCGGGAGAAAGTCTCTGAGGGCTTGGAAAGTGTACCGCGGACAC | 1158 |
| Db | 1081 | CACGTCTTCTTCGTGACCCCGGGAGAAAGTCTCTGAGGGCTTGGAAAGTGTACCGCGGACAC | 1140 |
| Qy | 1151 | AGCAACCTGCACCAAGTTGCTGAGGACCTTCATTTGCGAGATCATCATCAACAGCAAT | 1210 |
| Db | 1141 | AGCAACCTGCACCAAGTTGCTGAGGACCTTCATTTGCGAGATCATCATCAACAGCAAT | 1200 |
| Qy | 1211 | TACACCGATGAGAGAGCAACTATGACATGCGCCTCATGCGGCTGTCCAAAGCCCTGACC | 1270 |
| Db | 1201 | TACACCGATGAGAGAGCAACTATGACATGCGCCTCATGCGGCTGTCCAAAGCCCTGACC | 1260 |
| Qy | 1271 | CTGTCCGCTCAATCCACCCCTGCTTCCCTCCCAATGCAATGACACAGACTCTTAAGCTCAAT | 1338 |
| Db | 1261 | CTGTCCGCTCAATCCACCCCTGCTTCCCTCCCAATGCAATGACACAGACTCTTAAGCTCAAT | 1320 |
| Qy | 1331 | GAGACCTGCTGATATCACAAGCTTTTGGACAAGACAGAGGAGACAGATGACAAGACATCCCC | 1390 |
| Db | 1321 | GAGACCTGCTGATATCACAAGCTTTTGGACAAGACAGAGGAGACAGATGACAAGACATCCCC | 1380 |
| Qy | 1391 | TTCTCCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTCTGTGTC | 1450 |
| Db | 1381 | TTCTCCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTCTGTGTC | 1440 |
| Qy | 1451 | TATGACATTAATCTTAACCCCAAGATGATATGTGTCTGTGGGACCTTGTGTGGGGCAGAGAC | 1510 |
| Db | 1441 | TATGACATTAATCTTAACCCCAAGATGATATGTGTCTGTGGGACCTTGTGTGGGGCAGAGAC | 1500 |
| Qy | 1511 | TCTGTGACAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAGAACCAACCGCTGTGATCTG | 1570 |
| Db | 1501 | TCTGTGACAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAGAACCAACCGCTGTGATCTG | 1560 |
| Qy | 1571 | GCAGGTGTACACAGCTGGGAGCACAAGCTGTGGCCAGAGAAAACAACTGTGTGTACAC | 1630 |
| Db | 1561 | GCAGGTGTACACAGCTGGGAGCACAAGCTGTGGCCAGAGAAAACAACTGTGTGTACAC | 1620 |
| Qy | 1631 | AAAGTGACAGAAATTTCTTCCCTGTGATTTTACACAGAAATGAGAGC | 1675 |
| Db | 1621 | AAAGTGACAGAAATTTCTTCCCTGTGATTTTACACAGAAATGAGAGC | 1665 |

DE Human transmembrane serine protease (MTSP)-related cDNA #3.
XX
XX Human; transmembrane serine protease; MTSP; gene; ss;
KW cell surface protease; plasmin; prostate specific antigen; PSA;
KW proliferative disease; cell-surface protease-associated disease;
KW autoimmune disease; inflammatory disease; infectious disease;
KW endocrine disease; cancer; ocular disorder; cardiovascular disorder;
KW chronic inflammatory disease; wound; circulatory disorder;
KW dermatological disorder; restenosis; rheumatoid arthritis; psoriasis;
KW diabetic retinopathy; laser surgery; scarring;
KW glaucoma filtering surgery scarring; macular degeneration;
KW CRST syndrome; bacterial infection; viral disease; solid neoplasm;
KW vascular tumour; lung; colon; prostate; melanoma; Kaposi's sarcoma.
XX
OS Homo sapiens.
XX
XX US2004001801-A1.
XX
XX 01-JAN-2004.
XX
XX 23-MAY-2002; 2002US-00156214.
XX
XX 23-MAY-2002; 2002US-00156214.
XX
XX 23-MAY-2002; 2002US-00156214.
XX
XX (CORV-) CORVAS INT INC.
XX
XX
XX
XX Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Slov DV;
PI
XX
XX MPI; 2004-190126/18.
XX
XX P-PSDB; ADJ46919.
XX
XX
XX Conjugate useful for treating e.g. cancer, cell-surface protease-
PT associated diseases, comprising a peptidic substrate or nucleic acid
PT substrate linked to a therapeutic agent through a linker.
XX
XX
XX Claim 9; SEQ ID NO 25; 361pp; English.
XX
XX
XX The invention relates to a conjugate comprising a therapeutic agent and a
CC peptidic substrate or nucleic acid substrate linked to the agent
CC optionally through a linker or peptidic linker, where the peptidic
CC substrate is proteolytically cleaved by a cell surface protease or a
CC soluble, released or shed form conjugate to liberate the agent and the
CC conjugate is not substantially cleaved by plasmin or prostate specific
CC antigen (PSA). The conjugate is useful for treating a disease, which
CC involves administering a conjugate to a subject, where the disease is
CC preferably a proliferative disease or a cell-surface protease-associated
CC disease. The diseases include autoimmune diseases, inflammatory diseases,
CC infectious diseases, and endocrine diseases. The conjugate is useful for
CC treating a cell-surface protease-associated disease, which involves
CC administering a conjugate comprising an agent and a peptidic substrate to
CC a subject exhibiting symptoms of a cell-surface protease-associated
CC disorder, where the disease is selected from cancer, ocular disorders,
CC cardiovascular disorders, chronic inflammatory diseases, wounds,
CC circulatory disorders, dermatological disorders, restenosis, rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, scarring from laser
CC surgery, scarring from glaucoma filtering surgery, macular degeneration,
CC CRST syndrome, bacterial infections, viral diseases, solid neoplasms and
CC vascular tumours such as lung cancer, colon cancer, prostate cancer,
CC melanoma and Kaposi's sarcoma. This sequence represents cDNA encoding a
CC transmembrane serine protease (MTSP)-related polypeptide of the
CC invention.
XX
XX Sequence 2067 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;
XX

```
RESULT 12
ADJ46918
ID ADJ46918 standard; CDNA; 2067 BP
XX
AC ADJ46918;
XX
XX 06-MAY-2004 (first entry)
XX
```

| | Query Match | Similarity | Score | DB | Length |
|------------|-------------|--------------|-------------|------------|--------|
| Best Local | Similarity | 99.9% | Pred. No. 0 | | |
| Matches | 1664 | Conservative | 0 | Mismatches | 1 |
| | | | | Indels | 0 |
| | | | | Gaps | 0 |

| QY | 11 | ATGAGAGGGACACGCCACGGGAAATGCATCTCCAGCAAGAACCTTTCACTGAGGAGCATCT | 70 |
|----|----|---|-----|
| Db | 1 | ATGAGAGAGGGACACGCCACGGGAAATGCATCTCCAGCAAGAACCTTTCACTGAGGAGCATCT | 60 |
| QY | 71 | CCAGGCCACGAGCATCTCCAGTGGGACACCTCCAGGCCGGGACCTTCACGCCGAGGCATCT | 130 |

```
Db 61 CCAAGCCAGGATCTCCAGGCTGGGACAACCTCCAGCCGGGATCTCCAGCCAGGACATCT 120
Qy 131 CCAAGCCAGGATCTCCAGGCTGGGACAACCTCCAGCCGGGATCTCCAGCCAGGACATCT 190
Db 121 CCAAGCCAGGATCTCCAGGCTGGGACAACCTCCAGCCGGGATCTCCAGCCAGGACATCT 180
Qy 191 CCAAGCTGTACCTCCAGGCTGGGACAACCTCCAGCCGGGATCTCCAGCCAGGACATCT 250
Db 181 CCAAGCTGTACCTCCAGGCTGGGACAACCTCCAGCCGGGATCTCCAGCCAGGACATCT 240
Qy 251 CCAAGCCGGGATCTCCAGGCTGGGACAACCTCCAGCTTCAGGTCCTCATCCGAGGATCTCA 310
Db 241 CCAAGCCGGGATCTCCAGGCTGGGACAACCTTCAGGTCCTCATCCGAGGATCTCA 300
Qy 311 TCCGCGCAGGTCAGGCTGGGACAACCTCCCAACAGAGTACCTTTGAGACA 370
Db 301 TCCGCGCAGGTCAGGCTGGGACAACCTCCCAACAGAGTACCTTTGAGACA 360
Qy 371 CCAAGTGGGGGCTGATCCATCCGATCTCTCCAGTCTAGAGCAAGCAAGGAGCC 430
Db 361 CCAAGTGGGGGCTGATCCATCCGATCTCTCCAGTCTAGAGCAAGCAAGGAGCC 420
Qy 431 ACCAGGAGAGCCAGGATACAGGCTGCGCAAGTTCACTGGCGGAGAGCCAGAGAG 490
Db 421 ACCAGGAGAGCCAGGATACAGGCTGCGCAAGTTCACTGGCGGAGAGCCAGAGAG 480
Qy 491 CTACCGCTCATCGGGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 550
Db 481 CTACCGCTCATCGGGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Qy 551 TTCCAGTTCTGGGAGGAGCAACAGAGATCAGAGTCAAGAGAGAGAGAGAGAGTCC 610
Db 541 TTCCAGTTCTGGGAGGAGCAACAGAGATCAGAGTCAAGAGAGAGAGAGAGTCC 600
Qy 611 AAGCAGCCTGTTCCGCTGTGACGCGGCTGGTGTGACGTCAAGCTGAGAGTCAAGCTGG 670
Db 601 AAGCAGCCTGTTCCGCTGTGACGCGGCTGGTGTGACGTCAAGCTGAGAGTCAAGCTGG 660
Qy 671 TGCCTGAGGTTTGAAGGAGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 730
Db 661 TGCCTGAGGTTTGAAGGAGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Qy 731 TGGCTTCCATCTCTGAGAGCACTGGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 790
Db 721 TGGCTTCCATCTCTGAGAGCACTGGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Qy 791 CTGGGTTTCCAGAGTGTCTCAACCGGACAACGAGGTTGCCACAGAGATTTTGGCAACAG 850
Db 781 CTGGGTTTCCAGAGTGTCTCAACCGGACAACGAGGTTGCCACAGAGATTTTGGCAACAG 840
Qy 851 TTCTCAATCTTGAATACCACTCCACATCCAGGAAAGCTTCAAGTCTGAATGCGCT 910
Db 841 TTCTCAATCTTGAATACCACTCCACATCCAGGAAAGCTTCAAGTCTGAATGCGCT 900
Qy 911 TCCAGAGGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
Db 901 TCCAGAGGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Qy 971 ATCGTGGAGGAGGAGGCTGCGGCTCGGATGAGAAAGTGGCTTGGCAAGTGTGATCTG 1030
Db 961 ATCGTGGAGGAGGAGGCTGCGGCTCGGATGAGAAAGTGGCTTGGCAAGTGTGATCTG 1020
Qy 1031 GGCACCAACCACTCTGTGAGGAGCAGCTCATGAGCGCCAGTGGTGTCTCACTGCGCC 1090
Db 1021 GGCACCAACCACTCTGTGAGGAGCAGCTCATGAGCGCCAGTGGTGTCTCACTGCGCC 1080
Qy 1091 CACGCTTCTCTGAGACCGGAGAGAGGCTCTGAGAGGCTGGAAGGTGTAGCGCGGACCC 1150
Db 1081 CACGCTTCTCTGAGACCGGAGAGAGGCTCTGAGAGGCTGGAAGGTGTAGCGCGGACCC 1140
Qy 1151 AGCAACCTGACCAAGTGTGCTGAGGAGGCTTCATGCGAGATCATCATCAAGCAAT 1210
Db 1210 AGCAACCTGACCAAGTGTGCTGAGGAGGCTTCATGCGAGATCATCATCAAGCAAT 1200

Db 1141 AGCAACCTGACCAAGTGTGCTGAGGAGGCTTCATGCGAGATCATCATCAAGCAAT 1200
Qy 1211 TACACCGATGAGAGAGCACTATGATGAGCCCTCAATGGGCTGTCAAGCCCTGAC 1270
Db 1201 TACACCGATGAGAGAGCACTATGATGAGCCCTCAATGGGCTGTCAAGCCCTGAC 1260
Qy 1271 CTGTCCGCTCAATCCACCTGCTGCTGCTCCCATGATGAGAGACCTTTAGCTCAAT 1330
Db 1261 CTGTCCGCTCAATCCACCTGCTGCTGCTCCCATGATGAGAGACCTTTAGCTCAAT 1320
Qy 1331 GAGACTGTGATGATGAGGCTTTGGCAAGACAGGAGACAGATGACAGATGATCC 1390
Db 1321 GAGACTGTGATGATGAGGCTTTGGCAAGACAGGAGACAGATGACAGATGATCC 1380
Qy 1391 TTCTCCGAGGAGGATGAGGCTATGATGATGATGATGATGATGATGATGATGATG 1450
Db 1381 TTCTCCGAGGAGGATGAGGCTATGATGATGATGATGATGATGATGATGATGATG 1440
Qy 1451 TATGACAGTACCTTACCCCAAGATGATGATGATGATGATGATGATGATGATGATG 1510
Db 1441 TATGACAGTACCTTACCCCAAGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy 1511 TCTGCGCAGGAGAGCAGCGGGGCTCTTGTCTGTGAGAGAGCAACCGCTGTACTG 1570
Db 1501 TCTGCGCAGGAGAGCAGCGGGGCTCTTGTCTGTGAGAGAGCAACCGCTGTACTG 1560
Qy 1571 GCAGGTGTACACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAACTGTGTGTACAC 1630
Db 1561 GCAGGTGTACACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAACTGTGTGTACAC 1620
Qy 1631 AAGTGCAGAAAGTTCTTCCCTGATTTACAGCAAGATGAGAGAGC 1675
Db 1621 AAGTGCAGAAAGTTCTTCCCTGATTTACAGCAAGATGAGAGAGC 1665

RESULT 13
ADY50053
ID ADY50053 standard; DNA; 2067 BP.
XX
AC ADY50053;
XX
DT 19-MAY-2005 (first entry)
XX
DE Human endothelias-2L DNA.
XX
KW Antiarthritic; Antirheumatic; Antipsoriatic; Antidiabetic;
XX Ophthalmological; Cardiovascular-Gen.; Vulnery; Antinflammatory;
XX Vasoectopic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
XX rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
XX dermatological disease; cancer; neoplasm; endothelias-2; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2005019270-A2.
XX
PD 03-MAR-2005.
XX
PF 12-AUG-2004; 2004WO-US026148.
XX
PR 14-AUG-2003; 2003US-0495005P.
XX
PR 14-NOV-2003; 2003US-0520164P.
XX
PA (DYAX-) DYAX CORP.
XX PA (BEND-) DENDREON CORP.
XX PI Madison EL, Nixon A;
XX
XX WPI: 2005-202609/21.
XX P-PSDB: ADY50054.
XX
PT Novel protein capable of inhibiting endothelias-2, useful for treating
XX or preventing angiogenesis related disorder e.g. cancer.
XX
```

PS Disclosure; SEQ ID NO 1; 157bp; English.
CC The invention relates to an isolated protein (I) which comprises a heavy
CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
CC immunoglobulin variable domain sequence, where the first and second
CC immunoglobulin variable domain sequences from an antigen binding site
CC that specifically binds to human endothelase-2. (I) is useful for
CC detecting an endothelase or endothelase activity in a sample, for
CC modulating an activity of an E72-expressing cell, for modulating
CC proteolysis, for killing or inhibiting growth of a cell, for detecting
CC endothelase in a subject, for modulating endothelase activity in a
CC subject, for treating or preventing a disorder characterized by unwanted
CC angiogenesis in a subject. The disorder is chosen from rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, ocular disorder such as
CC pterygii recurrence, scarring excimer laser surgery and glaucoma
CC filtering surgery, cardiovascular disorders, chronic inflammatory
CC disorders, wound repair, circulatory disorders, crest syndromes,
CC dermatological disorders and cancers. The present sequence represents the
CC human endothelase-2L DNA.
XX
SQ Sequence 2067 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;
Query Match 95.2%; Score 1663.4; DB 14; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 11 ATGAGAGAGGACACACCGGAGATGATCTCTCAGCAAGAAACACCTTCAGCTGAGCATCT 70
DB 1 ATGAGAGAGGACACACCGGAGATGATCTCTCAGCAAGAAACACCTTCAGCTGAGCATCT 60
QY 71 CGAGCCCAAGGATCTCCAGCTGGAGACACCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 130
DB 61 CGAGCCCAAGGATCTCCAGCTGGAGACACCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 120
QY 131 CGAGCCCAAGGATCTCCAGCTGGAGACACCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 190
DB 121 CGAGCCCAAGGATCTCCAGCTGGAGACACCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 180
QY 191 CGAGCTGTACCTCCAGGCGCGGATCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 250
DB 181 CGAGCTGTACCTCCAGGCGCGGATCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 240
QY 251 CGAGCCCAAGGATCTCCAGCTGGAGACACCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 310
DB 241 CGAGCCCAAGGATCTCCAGCTGGAGACACCTTCAGGCGCGGATCTCCAGCCCAAGCATCT 300
QY 311 TCCGCCAGGTAGAGCTCGGTGAGCAACCTCCCAACAGATGATCTTGTATGAGCAACA 370
DB 301 TCCGCCAGGTAGAGCTCGGTGAGCAACCTCCCAACAGATGATCTTGTATGAGCAACA 360
QY 371 CGAGTGGGGGCTGATCCATCCATCTCTCTGAGGTGAGCAACAGCAAGAGGAGCC 430
DB 361 CGAGTGGGGGCTGATCCATCCATCTCTCTGAGGTGAGCAACAGCAAGAGGAGCC 420
QY 431 ACCAGGAGAGCCAGGATCGAGCTGCGCAAGTTCACTTGGCGGAGGCGCAAGAGAG 490
DB 421 ACCAGGAGAGCCAGGATCGAGCTGCGCAAGTTCACTTGGCGGAGGCGCAAGAGAG 480
QY 491 CTACCGCTCATCGGAGTGGTCTCTCTCTCATGCGCCCTGGGTGTTTGGCTATATCTTC 550
DB 481 CTACCGCTCATCGGAGTGGTCTCTCTCTCATGCGCCCTGGGTGTTTGGCTATATCTTC 540
QY 551 TTCCAGTTCTGGAGGAGCACAAGAGATCAGATCAAGAGAGAGAGAGAGTGTCTCC 610
DB 541 TTCCAGTTCTGGAGGAGCACAAGAGATCAGATCAAGAGAGAGAGAGAGTGTCTCC 600
QY 611 AAGCAGCGTGTTCGCTGTGACGCGGTGGTGAATGCAAGCTGAAGAGTGAAGAGTGGGC 670
DB 601 AAGCAGCGTGTTCGCTGTGACGCGGTGGTGAATGCAAGCTGAAGAGTGAAGAGTGGGC 660
QY 671 TGGGTGAGGTTTGAATGAGCAAGCTGTGTTAAATCTAATCTGAGGTCTCTCCATCAG 730
DB 661 TGGGTGAGGTTTGAATGAGCAAGCTGTGTTAAATCTAATCTGAGGTCTCTCCATCAG 720

QY 731 TGGCTTCCATCTGTAGACAGCAACTGGAATGACTCTTACTCAGAGAAAGACTGCGACAG 790
DB 721 TGGCTTCCATCTGTAGACAGCAACTGGAATGACTCTTACTCAGAGAAAGACTGCGACAG 780
QY 791 CTGGGTTTCAAGAGTGTCTCACCGGACCAACGAGGTTTCCCAAGGATTTTGGCAACAGC 850
DB 781 CTGGGTTTCAAGAGTGTCTCACCGGACCAACGAGGTTTCCCAAGGATTTTGGCAACAGC 840
QY 851 TTCTCAATCTTGAATGATCAACTCCACCATCATGAGAAAGCTCCACAGGTGAAATGCGCT 910
DB 841 TTCTCAATCTTGAATGATCAACTCCACCATCATGAGAAAGCTCCACAGGTGAAATGCGCT 900
QY 911 TCCAGGAGATATATCTCCCTCAAGTGTTCCTCACTGCGAGCTGAGAGGACATGACCGGCGG 970
DB 901 TCCAGGAGATATATCTCCCTCAAGTGTTCCTCACTGCGAGCTGAGAGGACATGACCGGCGG 960
QY 971 ATCTGTGAGAGGCGGCTGAGCTCGAATGACAGATGAGCTTGGCAAGTGAATCTGCACTTC 1030
DB 961 ATCTGTGAGAGGCGGCTGAGCTCGAATGACAGATGAGCTTGGCAAGTGAATCTGCACTTC 1020
QY 1031 GGCACCAACCAATCTGTGAGAGGAGCTCATTTGAGCGCCAGTGGGTGCTCACTGCGGC 1090
DB 1021 GGCACCAACCAATCTGTGAGAGGAGCTCATTTGAGCGCCAGTGGGTGCTCACTGCGGC 1080
QY 1091 CACTGCTTCTTGTGACCCCGGAGAAAGTCTGAGAGGCTGGAAGGTGTAACGCGGACCC 1150
DB 1081 CACTGCTTCTTGTGACCCCGGAGAAAGTCTGAGAGGCTGGAAGGTGTAACGCGGACCC 1140
QY 1151 AGCAACTGACCAAGTTGCTGAGGAGGAGCTCCATTCGAGATCATCAACAGCAAT 1210
DB 1141 AGCAACTGACCAAGTTGCTGAGGAGGAGCTCCATTCGAGATCATCAACAGCAAT 1200
QY 1211 TACACGATGAGAGAGAGAGCTATGATGATGAGCTTCAAGGAGCTGCAAGCCCTGACC 1270
DB 1201 TACACGATGAGAGAGAGAGCTATGATGATGAGCTTCAAGGAGCTGCAAGCCCTGACC 1260
QY 1271 CTGTCCGCTCAATCCACCTGCTTGGCTCCCATGAGTGAACAGACCTTTAGCTCAAT 1330
DB 1261 CTGTCCGCTCAATCCACCTGCTTGGCTCCCATGAGTGAACAGACCTTTAGCTCAAT 1320
QY 1331 GAGACTGTGATGATCAAGGCTTTGGCAAGACAGGAGACATGATGACATATCCCCC 1390
DB 1321 GAGACTGTGATGATCAAGGCTTTGGCAAGACAGGAGACATGATGACATATCCCCC 1380
QY 1391 TTCTCCGAGGAGGAGGAGTCAATCTCATGAGCTTCAAGAAATGCAATGACTACTTGTCTC 1450
DB 1381 TTCTCCGAGGAGGAGGAGTCAATCTCATGAGCTTCAAGAAATGCAATGACTACTTGTCTC 1440
QY 1451 TATGACATTAACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTGGGGGAGAGAC 1510
DB 1441 TATGACATTAACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTGGGGGAGAGAC 1500
QY 1511 TCTGTCCAGGAGAGACAGCGGGGGGCTCTTGTCTGTATGAGCAACACCGCTGTACTCTG 1570
DB 1501 TCTGTCCAGGAGAGACAGCGGGGGGCTCTTGTCTGTATGAGCAACACCGCTGTACTCTG 1560
QY 1571 GAGAGTGTCAACAGCTGAGGAGCAAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACACC 1630
DB 1561 GAGAGTGTCAACAGCTGAGGAGCAAGGCTGTGGCCAGAGAGAAACAAACCTGTGTGTACACC 1620
QY 1631 AAAGTGAACAGAGTTCTTCCCTGTGATTTACAGCAAGATGAGAGAC 1675
DB 1621 AAAGTGAACAGAGTTCTTCCCTGTGATTTACAGCAAGATGAGAGAC 1665
RESULT 14
ADH17413
ID ADH17413 standard; cDNA; 2432 BP.
XX
AC ADH17413;
XX
DT 11-MAR-2004 (first entry)

```

Db      181  CCAGGCAATCTCCAGTGTGACACCTCCAGCCGCGGCAATCTCCAGCCGCGGCAATCTCCAGC  240
Qy      241  CCAGGCAATCTCCAGCCGCGGCAATCTCCAGGCTCTGGCAATCACTTTCCAGGCTCTCAATCCG  300
Db      241  CCAGGCAATCTCCAGCCGCGGCAATCTCCAGGCTCTGGCAATCACTTTCCAGGCTCTCAATCCG  300
Qy      301  CAGGTCATCATCCGCGAGGTCAAGCTCTGGTGAACAACCTCCCAACAGAGTGAACCTTTGT  360
Db      301  CAGGTCATCATCCGCGAGGTCAAGCTCTGGTGAACAACCTCCCAACAGAGTGAACCTTTGT  360
Qy      361  TAGAGCAACACCAATGGGGGCTGTATCCCATCCAGTCAATCTCTGCGAGGTGACAGACCG  420
Db      361  TAGAGCAACACCAATGGGGGCTGTATCCCATCCAGTCAATCTCTGCGAGGTGACAGACCG  420
Qy      421  AACCAAGGCGCCACAGAGAGAGCCAGAGTACAGAGCTGCGCAAGTTCACTTGCGGAGAGG  480
Db      421  AACCAAGGCGCCACAGAGAGAGCCAGAGTACAGAGCTGCGCAAGTTCACTTGCGGAGAGG  480
Qy      481  CCAGAGAGCTACCGCTCATTCGGGTGCTGCTCTCTCTCAATTCGCTGTTGCTTCTGCT  540
Db      481  CCAGAGAGCTACCGCTCATTCGGGTGCTGCTCTCTCTCTCAATTCGCTGTTGCTTCTGCT  540
Qy      541  CATCATCTCTCTTCAGATTCCTGCGAGGCGCACAGAGGATCAAGATCAAGAGAGAGAGGA  600
Db      541  CATCATCTCTCTTCAGATTCCTGCGAGGCGCACAGAGGATCAAGATCAAGAGAGAGAGGA  600
Qy      601  GAGCTGTCCCAAGAGACGCTGTTCCGCTGTGAAGGGTGTGTGAAGTCAAGTCAAGAGAGTGA  660
Db      601  GAGCTGTCCCAAGAGACGCTGTTCCGCTGTGAAGGGTGTGTGAAGTCAAGTCAAGAGAGTGA  660
Qy      661  CGAGCTGGGCTGCGTGAAGTTGAATGAGGACAGTCTCTGCTTAAATCTAATCTGAGTCT  720
Db      661  CGAGCTGGGCTGCGTGAAGTTGAATGAGGACAGTCTCTGCTTAAATCTAATCTGAGTCT  720
Qy      721  CTCCCATCAGTGGCTTCCCATCTGTAGAGAGCACTGGAATACCTCTCTCTCTCTCTCTCT  780
Db      721  CTCCCATCAGTGGCTTCCCATCTGTAGAGAGCACTGGAATACCTCTCTCTCTCTCTCTCT  780
Qy      781  CTGCGAGAGCTGGGTTTCTGAGAGTGTCTACCGGACCAACCGAGGTTGCCCAAGAGATTT  840
Db      781  CTGCGAGAGCTGGGTTTCTGAGAGTGTCTACCGGACCAACCGAGGTTGCCCAAGAGATTT  840
Qy      841  TGCCCAAGCTTCTCAATCTTGAATACACTCCACATCCAGAGAAAGCTTCCACAGGTC  900
Db      841  TGCCCAAGCTTCTCAATCTTGAATACACTCCACATCCAGAGAAAGCTTCCACAGGTC  900
Qy      901  TGAATGCCCTTCCAGAGGTAATCTCTCTCAAGTGTCCCACTGCGGACTGAGGGCCAT  960
Db      901  TGAATGCCCTTCCAGAGGTAATCTCTCTCAAGTGTCCCACTGCGGACTGAGGGCCAT  960
Qy      961  GACCGGCGAGTCTGTGAGAGGGGCGCTGCTCGGATAGCAAGTGGCTTGGCAAGTGAG  1020
Db      961  GACCGGCGAGTCTGTGAGAGGGGCGCTGCTCGGATAGCAAGTGGCTTGGCAAGTGAG  1020
Qy      1021  TCTGCACTTGGGACCAACCAATCTGTGAGAGCAAGCTCAATTAAGCGCCAGTGGTCT  1080
Db      1021  TCTGCACTTGGGACCAACCAATCTGTGAGAGCAAGCTCAATTAAGCGCCAGTGGTCT  1080
Qy      1081  CACTGCGGCGGCACTGCTTCTTCTGTAACCGGGAGAAAGTCTTGAAGGGCTGAAAGTGA  1140
Db      1081  CACTGCGGCGGCACTGCTTCTTCTGTAACCGGGAGAAAGTCTTGAAGGGCTGAAAGTGA  1140
Qy      1141  CGCGGGGCGAGCAACCTGCAACAGTTGCTGAGGAGCTCTCAATTCGCGAGATCATAT  1200
Db      1141  CGCGGGGCGAGCAACCTGCAACAGTTGCTGAGGAGCTCTCAATTCGCGAGATCATAT  1200
Qy      1201  CAAACAGCAATTACACCGATGAGAGAGAGCACTATGACATGCGCTCATGCGGCTGTCAA  1260
Db      1201  CAAACAGCAATTACACCGATGAGAGAGAGCACTATGACATGCGCTCATGCGGCTGTCAA  1260
Qy      1261  GCGCCTGACCTGTTCGGTGTCAATCCAGCTGTTGCTCTCCCATGATGACAGAACTT  1320
Db      1261  GCGCCTGACCTGTTCGGTGTCAATCCAGCTGTTGCTCTCCCATGATGACAGAACTT  1320

```

```

Qy      1321  TAGCCTCAATGAGACTGTGTGATCAAGGCTTTGGCAAGACAGAGGAGACAGATGACAA  1380
Db      1321  TAGCCTCAATGAGACTGTGTGATCAAGGCTTTGGCAAGACAGAGGAGACAGATGACAA  1380
Qy      1381  GACATCCCGCTTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGA  1440
Db      1381  GACATCCCGCTTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGA  1440
Qy      1441  CTACTGGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG  1500
Db      1441  CTACTGGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG  1500
Qy      1501  GGGCAGAGACTCTGCCAGGAGACAGCGGGGGGCTTTGTCTGTGACAGAAACAACG  1560
Db      1501  GGGCAGAGACTCTGCCAGGAGACAGCGGGGGGCTTTGTCTGTGACAGAAACAACG  1560
Qy      1561  CTGCTACCTGGCAGAGTGTCAACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACCTGG  1620
Db      1561  CTGCTACCTGGCAGAGTGTCAACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACCTGG  1620
Qy      1621  TGTGTACACCAAGTGAAGAAAGTTCTTCCCTGATTTACAGAAAGATGAGAGAGAGT  1680
Db      1621  TGTGTACACCAAGTGAAGAAAGTTCTTCCCTGATTTACAGAAAGATGAGAGAGAGT  1680
Qy      1681  GCGATTGAGAAATCTTAACCAAGCTGGCTGTGCTGTGCAACAGACCGGCTGTGTGAC  1740
Db      1681  GCGATTGAGAAATCTTAACCAAGCTGGCTGTGCTGTGCAACAGACCGGCTGTGTGAC  1740
Qy      1741  TCGAGAAA  1748
Db      1741  TCGAGAAA  1748

RESULT 2
US-10-177-661-1
; Sequence 1, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Vircea, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177,661
; PRIOR FILING DATE: 2002-06-20
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1434)
; OTHER INFORMATION:
US-10-177-661-1

Query Match      72.0%; Score 1258.6; DB 3; Length 1434;
Best Local Similarity 98.5%; Pred. No. 2.2e-237;
Matches 1270; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Db 266 TGGTTGCTCATCTCTTCCAGTTCTGGACGGGCAACAGGGATCAGTACAGG 325
 QY 591 AGCAGAGGAGAGCTGTCTCCCAAGCAGCTGTCTGTCAGGGGTGTGACTGCAAGC 650
 Db 326 AGCAGAGGAGAGCTGTCTCCCAAGCAGCTGTCTGTCAGGGGTGTGACTGCAAGC 385
 QY 651 TGAAGAGTGAAGAGCTGGGCTGTGAGGTTGACTGGGCAAGTCTCTGCTTAAATCT 710
 Db 386 TGAAGAGTGAAGAGCTGGGCTGTGAGGTTGACTGGGCAAGTCTCTGCTTAAATCT 445
 QY 711 ACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAAATGACTCTACT 770
 Db 446 ACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAAATGACTCTACT 505
 QY 771 CAGAGAAAGCTGTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACAGAGTTGGCC 830
 Db 506 CAGAGAAAGCTGTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACAGAGTTGGCC 565
 QY 831 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCC 890
 Db 566 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCC 625
 QY 891 TCCACAGGTCTGAATGCGCTTCCAGCGGTATATCTCCCTCAGTGTTCCTCAGTGGCAGC 950
 Db 626 TCCACAGGTCTGAATGCGCTTCCAGCGGTATATCTCCCTCAGTGTTCCTCAGTGGCAGC 685
 QY 951 TGAAGGCGCATGACCGGGCGGATCTGTGGAGGGGCGCTGGCTCCGATAGCAAGTGGCTTT 1010
 Db 686 TGAAGGCGCATGACCGGGCGGATCTGTGGAGGGGCGCTGGCTCCGATAGCAAGTGGCTTT 745
 QY 1011 GGCAGAGTGAAGTGCATCTTGGGCAACCCACATCTGTGGAGGACGCTCATTTGACGCGC 1070
 Db 746 GGCAGAGTGAAGTGCATCTTGGGCAACCCACATCTGTGGAGGACGCTCATTTGACGCGC 805
 QY 1071 AGTGGGTGCTCACTGCGGCCCACTGCTTCTTCTGTGACCCCGGAGAGAGTCTCTGAGGGCT 1130
 Db 806 AGTGGGTGCTCACTGCGGCCCACTGCTTCTTCTGTGACCCCGGAGAGAGTCTCTGAGGGCT 865
 QY 1131 GGAAGGTGACCGGGCGGCAACAGCACTGTGACAGTGTGCTGAGGACGCTTCTCATTTGCGG 1190
 Db 866 GGAAGGTGACCGGGCGGCAACAGCACTGTGACAGTGTGCTGAGGACGCTTCTCATTTGCGG 925
 QY 1191 AGATCATCATCAAGCAATTAACAACGATGAGAGAGCACTATGACATGCGCCCTCATGCG 1250
 Db 926 AGATCATCATCAAGCAATTAACAACGATGAGAGAGCACTATGACATGCGCCCTCATGCG 985
 QY 1251 GGCCTGTCCAAAGCCCTGACCTGTCCGCTCAATCCACCTGTGCTTCCCTCCCATGCAATG 1310
 Db 986 GGCCTGTCCAAAGCCCTGACCTGTCCGCTCAATCCACCTGTGCTTCCCTCCCATGCAATG 1045
 QY 1311 GACAGACCTTTAGCTCTCAATGAGACCTGTGATCAAGGCTTTGGCAAGCAAGGAGGA 1370
 Db 1046 GACAGACCTTTAGCTCTCAATGAGACCTGTGATCAAGGCTTTGGCAAGCAAGGAGGA 1105
 QY 1371 CAGATGACAAACATCCCTTCTCCGCGGAGGTGACAGGTCAATCTCATGCACTTCAAGA 1430
 Db 1106 CAGATGACAAACATCCCTTCTCCGCGGAGGTGACAGGTCAATCTCATGCACTTCAAGA 1165
 QY 1431 AATCAATGACTACTTGTGTATGACAGTTACTTACCCCAAGAGATGATGTGTGTGGGG 1490
 Db 1166 AATCAATGACTACTTGTGTATGACAGTTACTTACCCCAAGAGATGATGTGTGTGGGG 1225
 QY 1491 ACCTTGTGGGGGCAAGACTTCTGCGGAGAGACAGCGGGGGGCTTTGTCTGTGAGC 1550
 Db 1226 ACCTTGTGGGGGCAAGACTTCTGCGGAGAGACAGCGGGGGGCTTTGTCTGTGAGC 1285
 QY 1551 AGAACAACCGGTGATCTGAGAGGTGACAGAGTGGGGCAAGGCTGTGGCCAGAGAA 1610
 Db 1286 AGAACAACCGGTGATCTGAGAGGTGACAGAGTGGGGCAAGGCTGTGGCCAGAGAA 1345
 QY 1611 ACAAACTGTGTGTACACAAAGTGAAGAGTTCTTCTCTGATTTTACAGCAAGATGG 1670
 Db 1346 ACAAACTGTGTGTACACAAAGTGAAGAGTTCTTCTCTGATTTTACAGCAAGATGG 1405

QY 1671 AGAGCGAGTGCATTCAGAAAAATCTTAA 1699
 Db 1406 AGAGCGAGTGCATTCAGAAAAATCTTAA 1434
 RESULT 3
 US-10-177-661-3
 : Sequence 3, Application US/10177661
 : Patent No. 6794173
 : GENERAL INFORMATION:
 : APPLICANT: Anderson, Dirk M.
 : APPLICANT: Virca, G. Duke
 : TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
 : FILE REFERENCE: 3256-A
 : CURRENT APPLICATION NUMBER: US/10/177,661
 : PRIOR FILING DATE: 2002-06-20
 : PRIOR APPLICATION NUMBER: US 60/299,606
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 3
 : LENGTH: 1341
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)..(1341)
 : OTHER INFORMATION:
 : US-10-177-661-3
 Query Match 60.8%; Score 1062.6; DB 3; Length 1341;
 Best Local Similarity 91.3%; Pred. No. 1.5e-249;
 Matches 1177; Conservative 0; Mismatches 19; Indels 93; Gaps 1;
 QY 411 CAGCAGCAGCAACCAAGGCGCCACAGGAGAGCCAGGATCAGAGCTGCCAAGTTCACT 470
 Db 146 CAGCCTCATATATCCCTGCTGCGGCCCTGCTCCTAGGATGAGAGCTGCCAAGTTCACT 205
 QY 471 GCGGAGGCGCAGAGCAGCTACCGCTCATCGGGTGTGCTCTCTCATTTGCCCTGG 530
 Db 206 GCGGAGGCGCAGAGCAGCTACCGCTCATCGGGTGTGCTCTCTCATTTGCCCTGG 265
 QY 531 TGGTTTGGCTCATCATCTCTTCCAGTTCTGGGAGGCGCACAGGGATCAGGTACAAG 590
 Db 266 TGGTTTGGCTCATCATCTCTTCCAGTTCTGGGAGGCGCACAGGTCTCTCTCATTTGCCCTGG 325
 QY 591 AGCAGAGGAGAGCTGTCCCAAGCAGCTGTCTGCTGTGACGGGGTGTGACTGCAAGC 650
 Db 326 AGCAGAGGAGAGCTGTCCCAAGCAGCTGTCTGCTGTGACGGGGTGTGACTGCAAGC 385
 QY 651 TGAAGAGTGAAGAGCTGGGCTGTGAGGTTGACTGGAGACAGTCTCTGTTAAATCT 710
 Db 386 TGAAGAGTGAAGAGCTGGGCTGTGAGGTTGACTGGAGACAGTCTCTGTTAAATCT 445
 QY 711 ACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAAATGACTCTACT 770
 Db 446 ACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAAATGACTCTACT 505
 QY 771 CAGAGAAAGCTGTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACAGAGTTGGCC 830
 Db 506 CAGAGAAAGCTGTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACAGAGTTGGCC 537
 QY 831 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCC 890
 Db 538 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCC 537
 QY 891 TCCACAGGTCTGAATGCGCTTCCAGCGGTATATCTCCCTCAGTGTTCCTCAGTGGCAGC 950
 Db 538 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCC 537
 QY 951 TGAAGGCGCATGACCGGGCGGATCTGTGGAGGGGCGCTGCGGATAGCAAGTGGCTTT 1010

```
Db      593  TGAGGGCCATGACCGGGCGATCGTGAGAGGGGCGCTGCTCGATGACAAAGTGCCCTT 652
Qy      1011  GGCAAGTAGTCTGCACTTGGGCAACCACTCTGAGAGGACGCTCATTTAGCC 1070
Db      653  GGCAAGTAGTCTGCACTTGGGCAACCACTCTGAGAGGACGCTCATTTAGCC 712
Qy      1071  AGTGGGTGCTCACTGCGGCCCACTGCTTCTTGATGACCCGGGAGAAAGTCTTGAGGGCT 1130
Db      713  AGTGGGTGCTCACTGCGGCCCACTGCTTCTTGATGACCCGGGAGAAAGTCTTGAGGGCT 772
Qy      1131  GGAAGGTGTACCGGGGACCAAGCACTTGACCAAGTCTTGAGGACGCTTCATTCG 1190
Db      773  GGAAGGTGTACCGGGGACCAAGCACTTGACCAAGTCTTGAGGACGCTTCATTCG 832
Qy      1191  AGATCATCATCAAGCAATTAACACCGATGAGAGAGCACTATGACATCGCCCTCATGC 1250
Db      833  AGATCATCATCAAGCAATTAACACCGATGAGAGAGCACTATGACATCGCCCTCATGC 892
Qy      1251  GGTGTTCAGAGCCCTGACCTGTGCTGCAATCCACCTGCTTGCTCCCATGATG 1310
Db      893  GGTGTTCAGAGCCCTGACCTGTGCTGCAATCCACCTGCTTGCTCCCATGATG 952
Qy      1311  GACGACCTTTAGCTCAATGAGACCTGTGAGAGTGCAGAGCTTTGGCAAGACAGGAGA 1370
Db      953  GACGACCTTTAGCTCAATGAGACCTGTGAGAGTGCAGAGCTTTGGCAAGACAGGAGA 1012
Qy      1371  CAGATGACAAAGACATCCCTTCTCCGGAGGTGCAGAGTCAATCTCATGCACTTCAAGA 1430
Db      1013  CAGATGACAAAGACATCCCTTCTCCGGAGGTGCAGAGTCAATCTCATGCACTTCAAGA 1072
Qy      1431  AATCAATGACTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGG 1490
Db      1073  AATCAATGACTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGG 1132
Qy      1491  ACCTTGTGGGGGAGAGACCTCTGCGAGAGAGACAGGGGGGGCTTTGTCTGTAGC 1550
Db      1133  ACCTTGTGGGGGAGAGACCTCTGCGAGAGAGACAGGGGGGGCTTTGTCTGTAGC 1192
Qy      1551  AGAACAACCGCTGTACCTGTGACAGGTGTCACAGCTGGGGGCAAGGCTGTGGCAGAGAA 1610
Db      1193  AGAACAACCGCTGTACCTGTGACAGGTGTCACAGCTGGGGGCAAGGCTGTGGCAGAGAA 1252
Qy      1611  ACAAACTGTGTGTACCAACAAGTGAAGAAATTTCCCTGTGATTTTACAGCAAGATGG 1670
Db      1253  ACAAACTGTGTGTACCAACAAGTGAAGAAATTTCCCTGTGATTTTACAGCAAGATGG 1312
Qy      1671  AGAGCGAGTGCAGTTTCAGAAAATCTTAA 1699
Db      1313  AGAGCGAGTGCAGTTTCAGAAAATCTTAA 1341

RESULT 4
US-09-879-792-35
; Sequence 35, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xieo, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
(Docket No. 6734006 L10-81-WO)
```

```
; SEQ ID NO 35
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1230)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-792-35

Query Match      60.2%; Score 1053; DB 3; Length 1230;
Beet Local Similarity 92.1%; Pred. No. 3.3e-247;
Matches 1134; Conservative 0; Mismatches 84; Indels 13; Gaps 2;

Qy      521  ATTGCGCTGGGTGGTTTGGCTCATCATCTCTTCCAGTTTGGCAGGGCCACAGAGGATC 580
Db      1  ATGACCCAGCTGTGTGCTTCTTTCTCTAGTCCAGTTTGGCAGNNCCACAGANNATC 60
Qy      581  AGGTACAAAGAGCAGAGAGAGAGCTGTCCCAAGACGCTGTGCTGACGAGGTGGTG 640
Db      61  AGGTACAAAGAGCAGAGAGAGAGCTGTCCCAAGACGCTGTGCTGACGAGGTGGTG 120
Qy      641  GACTGCAAGCTGGAAGATGAGAGTGGCTGGCTGAGTTTGACTGGAGCAACTCTCTG 700
Db      121  GACTGCAAGCTGGAAGATGAGAGTGGCTGGCTGAGTTTGACTGGAGCAAGTCTCTG 180
Qy      701  CTTAAATCTACTGTGGGTCTCTCCATCAGTGGCTTCCATCTGTAGCAGCACTGGAAAT 760
Db      181  CTTAAATCTACTGTGGGTCTCTCCATCAGTGGCTTCCATCTGTAGCAGCACTGGAAAT 240
Qy      761  GACTCTACTCAGAGAAAGACCTGCGACAGCTGGGTTTGAAGTGTCTCACCGGACAAC 820
Db      241  GACTCTACTCAGAGAAAGACCTGCGACAGCTGGGTTTGAAGTGTCTCACCGGACAAC 300
Qy      821  GAGGTTCGCCACAGAGATTTTGGCAACAGCTTTCATCTTTGAGATCAATCCACATC 880
Db      301  GAGGTTCGCCACAGAGATTTTGGCAACAGCTTTCATCTTTGAGATCAATCCACATC 360
Qy      881  CAGGAAGAGCTCCACAGAGTGTGAATGCTTCCAGAGGATATATCTCTCCAGATGTTCC 940
Db      361  CAGGAAGAGCTCCACAGAGTGTGAATGCTTCCAGAGGATATATCTCTCCAGATGTTCC 420
Qy      941  CACTGCGAGCTGAGAGGACATGACCGGGCGAGTCGTGGAGAGGGGCTGGCTCGGATAGC 1000
Db      421  CACTGCGAGCTGAGAGGACATGACCGGGCGAGTCGTGGAGAGGGGCTGGCTCGGATAGC 480
Qy      1001  AAGTGGCTTGGCAAGTGTGCACTTGTGGACCAACCAATCTGTGAGAGGACGCTC 1060
Db      481  AAGTGGCTTGGCAAGTGTGCACTTGTGGACCAACCAATCTGTGAGAGGACGCTC 540
Qy      1061  ATTGACGCCAGTGGGTGTCTCATCTGCCGCCCACTGCTTCTTGATGACCCGGAGAAAGTTC 1120
Db      541  ATTGACGCCAGTGGGTGTCTCATCTGCCGCCCACTGCTTCTTGATGACCCGGAGAAAGTTC 600
Qy      1121  CTGAGAGGCTGGAAGGTGTACGCGGGGACCAAGCACTGCAACAGTTGCTTGAGGACGCC 1180
Db      601  CTGAGAGGCTGGAAGGTGTACGCGGGGACCAAGCACTGCAACAGTTGCTTGAGGACGCC 660
Qy      1181  TCCATTGCCAGATCATCATCAACAGCAATTTACCCGATGAGAGGAGCACTATGATCATC 1240
Db      661  TCCATTGCCAGATCATCATCAACAGCAATTTACCCGATGAGAGGAGCACTATGATCATC 720
Qy      1241  GCCCTCATGCGGCTGTCTCAAGCCCTGACCTGCTGCTCACTCAACCTGCTTGGCTC 1300
Db      721  GCCCTCATGCGGCTGTCTCAAGCCCTGACCTGCTGCTGCTGAGGAATCTGTGACTCCCGC 780
Qy      1301  CCAATGATGAGACAGACCTTAAAGCTTCAATGAGACCTGCTGATCAGAGCTTTGGCAA- 1359
Db      781  TCTCTGCCCCCGAGCCCGACAGACCTCTGACGCTTGCAGCTTGTGACAGATCTGTCAAC 840
Qy      1360  -----GACGAGGAGACAGAGACAGAGCAATCCCTTCCCTCCGGAGAGTCCAG 1408
Db      841  TCATATCCGGGCCCCCAAGCTTGTGACAGACAGACATCCCTTCTCCGGAGAGTCCAG 900
```

Best Local Similarity 92.7%; Free: No.
Matches 516; Conservative 0; Mismatch

| | | | | | | | |
|----------------------|-------|--------------|-------------|------------|-----|--------|-----|
| Query Match | 25.2% | Score | 440.8 | DB | 3 | Length | 678 |
| Best Local Similarly | 82.7% | Pred. | No. 7.4e-98 | | | | |
| Matches | 516 | Conservative | 0 | Mismatches | 107 | Indels | 1 |
| | | | | | | Gaps | 1 |

QY 320 TCAGCTCGTAGAACCTTCCCAACAGAGTGTACCTTTAGAGCAACAGAGTGGG 379
Db 1 TCAGCTCGTAGAACCTTCCCAACAGAGTGTACCTTTAGAGCAACAGAGTGGG 60
QY 380 GCTGTATCCATCCGATCATCTCTGCGAGGTGACGACCAAGGCGCAACAGAGGAG 439
Db 61 GCTGTATCCATCCGAGCATCTCTGCGAGGTGACGACCAAGGCGCAACAGAGGAG 120
QY 440 AGCCAGGTACGAGGCTGCGCAAGTTCAGTGGCGGAGGCGCAAGAGTACCTGCTC 499
Db 121 AGCCAGGTACGAGTTCCTCCCAAGTTCCTGCGAGGAGGCGCAAGAGTACCTGCTC 180
QY 500 ATCGGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 559
Db 181 ATCGGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 560 TGGCAGGCGCACACAGGAGTACAGTACAGAGGAGGAGGAGTGTCCAGAGAGCT 619
Db 241 TGGCAGGCGCACACAGGAGTACAGTACAGAGGAGGAGGAGTGTCCAGAGAGCT 300
QY 620 GTTGGCTGTACGCGGCTGTGGAAGTGAAGGAGTGAAGAGTGGGCTGGGAGAG 679
Db 301 GTTGGCTGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 360
QY 680 TTTGACTGGGACAAGTCTCTGCTTAAATCTACTCTGCTGCTGCTGCTGCTGCTGCT 739
Db 361 TTTGACTGGGACAAGTCTCTGCTTAAATCTACTCTGCTGCTGCTGCTGCTGCTGCT 420
QY 740 ATCTGTAGCAGCACTGGAATGACTCTTACTGAGAGAGAGCTGCGAGAGTGGTTTC 799
Db 421 GTCTGACGACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 800 GAGAGTGTCAACCGGACCAAGGTTGCGCAAGGAGTTTGGCAAGGCTTCAATC 859
Db 481 GAGAGTGTCAACCGGACCAAGGTTGCGCAAGGAGTTTGGCAAGGCTTCAATC 540
QY 860 TTGAGATCAACTCCACCATCCAGAGAGCTTCCACAGTGTGATGCTTCCACAGG 919
Db 541 TCCGATTCACACACACCATCCAGAGAGCTTCCACAGTGTGATGCTTCCACAGG 599
QY 920 TATATCTCTCCACAGTGTCCAC 943
Db 600 TATATCTCTCCACAGTGTCCAC 623

RESULT 7
US-09-879-792-33
Sequence 33, Application US/09879792
Patent No. 6734006

GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Transmembrane Seline
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-WO)
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 614
TYPE: DNA
ORGANISM: mouse
US-09-879-792-33

Query Match 23.5%; Score 411.6; DB 3; Length 614;
Best Local Similarity 90.1%; Pred. No. 9.6e-91;
Matches 463; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

QY 1189 CGAGATCATCATCAACAGCAATTAACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248
Db 1 CCAGATCATCATCAACAGCAATTAACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 1249 GCGGCTGTCCAAAGCCCTGAGCCCTGCGAGTCAATCAATCAATCAATCAATCAATCA 1308
Db 61 CAGGCTGTCCAAAGCCCTGAGCCCTGCGAGTCAATCAATCAATCAATCAATCAATCA 120
QY 1309 TGGCAGAGCTTTAGCTTCAATGAGAGCTGTGATGAGAGAGAGAGAGAGAGAGAGAG 1368
Db 121 CGGTCAAGCTTGGGCTCAATGAGAGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 178
QY 1369 GACAGATGACAGAGATCCCTCTCTCGGAGAGTGAAGTCAATCATCACTTCA 1428
Db 179 GACAGATGACAGAGATCCCTCTCTCGGAGAGTGAAGTCAATCATCACTTCA 238
QY 1429 GAAATGCAATGACTTGTGTATGACAGTTACCTTACCCCAAGAGATGATGTGCTGG 1488
Db 239 GAAATGCAATGACTTGTGTATGACAGTTACCTTACCCCAAGAGATGATGTGCTGG 298
QY 1489 GGAATCTGTGGGCGAGAGAGCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1548
Db 299 GGAATCTGTGGGCGAGAGAGCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
QY 1549 GCAGAGAACCGCTGTGATCTGCGAGAGTGTCAACAGTGTGGGCGAGAGTGTGGCGAG 1608
Db 359 GCAGAGAACCGCTGTGATCTGCGAGAGTGTCAACAGTGTGGGCGAGAGTGTGGCGAG 418
QY 1609 AAAAAGCTGTGTGTATGACCAAGAGTGAAGAGTGTCTTCCCTGATTTTACAGCAAGAT 1668
Db 419 AAAAAGCTGTGTGTATGACCAAGAGTGAAGAGTGTCTTCCCTGATTTTACAGCAAGAT 478
QY 1669 GGAGAGGAGGTGCAATTCAGAAAATCTTAACA 1702
Db 479 GGAGAGGAGGTGCAATTCAGAAAATCTTAACA 512

RESULT 8
US-09-879-792-1
Sequence 1, Application US/09879792
Patent No. 6734006

GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Transmembrane Seline
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-WO)
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 402
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) - (402)
OTHER INFORMATION: n = A,T,C or G
US-09-879-792-1

Query Match 19.9%; Score 348; DB 3; Length 402;
Best Local Similarity 97.8%; Pred. No. 2.8e-75;
Matches 362; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 903 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCACTGCGAGCTGAGGGCCATGA 962
DB 1 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCACTGCGAGCTGAGGGCCATGA 60

QY 963 CCGGGCGGATGTGGAGGGGGCGCTGCGCTGGATAGCAAGTGGCTTTGGCAAGTGA 1022
DB 61 CCGGGCGGATGTGGAGGGGGCGCTGCGCTGGATAGCAAGTGGCTTTGGCAAGTGA 120

QY 1023 TGCACCTTGGGACCAACCCACATCTGTGAGAGCGCTATGAGCCCGAGGGGGCTCA 1082
DB 121 TGCACCTTGGGACCAACCCACATCTGTGAGAGCGCTATGAGCCCGAGGGGGCTCA 180

QY 1083 CTGCGCGCCCATGTCTTCTGTG-AACCGGAGAAAGTCTGAGAGGGCTGAAAGTGTAC 1141
DB 181 CTGCGCGCCCATGTCTTCTGTG-AACCGGAGAAAGTCTGAGAGGGCTGAAAGTGTAC 240

QY 1142 GCGGGCACCAACCTGACCACTGCTTGAAGCGCTTCCTCCATTCGCGAGATCATC 1201
DB 241 GCGGGCACCAACCTGACCACTGCTTGAAGCGCTTCCTCCATTCGCGAGATCATC 300

QY 1202 AACGCAATTACCGGATGAGAGAGAGCACTATGACATCGCCCTCATGGGGCTGCCAAG 1261
DB 301 AACGCAATTACCGGATGAGAGAGAGCACTATGACATCGCCCTCATGGGGCTGCCAAG 360

QY 1262 CCCCTGACCC 1271
DB 361 NCCCTTGACC 370

RESULT 9
US-09-879-792-29
Sequence 29, Application US/09879792
Patent No. 6734006
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
APPLICANT: Gedrich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
TITLE OF INVENTION: Protease
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT
PRIOR FILING DATE: 2001-06-12 (Docket No. 6734006 LIO-81-WO)
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 621
TYPE: DNA
ORGANISM: mouse
US-09-879-792-29

Query Match 19.2%; Score 335.8; DB 3; Length 621;
Best Local Similarity 80.8%; Pred. No. 3.1e-72;
Matches 489; Conservative 0; Mismatches 107; Indels 9; Gaps 8;

QY 422 ACCAGGGCACCCAGAG-AGAGCCAGGATAGAGCGCTGCCAAGTTCACTGG-CCGAGAG 479
DB 121 ACCAGGGCACCCAGGATAGAGCGCTGCCAAGTTCACTGG-CCGAGAG 180

QY 480 GCCAAGACAGCTACCGCTCATCGGATGCGTCTCTCTCATTTGCCCTGGTGGTTTGGC 539
DB 181 CCCAAGACAGCTACCGCTCATCGGATGCGTCTCTCTCATTTGCCCTGGTGGTTTGGC 240

QY 540 TCATCATCTCTTCCAGTTCTGG-CAGGGCCACACAGGATCAGTACACAGACAGAG 598
DB 241 TCATCATCTCTTCCAGTTCTGG-CAGGGCCACACAGGATCAGTACACAGACAGAG 300

QY 599 GAGAGCTGTCCCAAGACAGCGCTTCCGCTGATGAGGGGGGGTGGAG-TCGAAGTGAAG 657
DB 301 GAGAGCTGTCCCAAGACAGCGCTTCCGCTGATGAGGGGGGGTGGAG-TCGAAGTGAAG 360

QY 658 TGACAGCTGGCTGCGTGAAGTTGACTGAGACAGAGTCTCTTAAATCTACTTGG 717
DB 361 TGACAGCTGGCTGCGTGAAGTTGACTGAGACAGAGTCTCTTAAATCTACTTGG 420

QY 718 GTCTTC-CCATCATGCTTCCATCTGTAGACAGCACTGAAATGATCTTACTGAGA 776
DB 421 GTCTTC-CCATCATGCTTCCATCTGTAGACAGCACTGAAATGATCTTACTGAGA 479

QY 777 AGACCTGCGAGC-AGCTGGGTTTGGAGAGTCTACCGGACACCGAGGTTGCCACAG 835
DB 480 AGACCTGCGAGC-AGCTGGGTTTGGAGAGTCTACCGGACACCGAGGTTGCCACAG 539

QY 836 GATTTTGCACACAGCTTCTCATCTTGAATACAACTCCACATCCAGAGAAAGCTTCCAC 895
DB 540 GATTTTGCACACAGCTTCTCATCTTGAATACAACTCCACATCCAGAGAAAGCTTCCAC 597

QY 896 AGGTC 900
DB 598 AGGTC 602

RESULT 10
US-09-280-116-22
Sequence 22, Application US/09280116A
Patent No. 6331427
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 610
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: trypsin-like serine proteases
NAME/KEY: misc.feature
LOCATION: (1)..(610)
OTHER INFORMATION: n = a, t, c, or g
US-09-280-116-22

Query Match 19.1%; Score 333.2; DB 3; Length 610;
Best Local Similarity 97.0%; Pred. No. 1.3e-71;
Matches 360; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 903 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCACTGCGAGCTGAGGGCCATGA 962
DB 1 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCACTGCGAGCTGAGGGCCATGA 60

QY 963 CCGGGCGGATGTGGAGGGGGCGCTGCGCTGGATAGCAAGTGGCTTTGGCAAGTGA 1022
DB 61 CCGGGCGGATGTGGAGGGGGCGCTGCGCTGGATAGCAAGTGGCTTTGGCAAGTGA 120

PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-WO)
 PRIOR FILING DATE: 2001-06-12
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 34
 LENGTH: 751
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-879-792-34

Query Match 15.2%; Score 265.4; DB 3; Length 751;
 Best Local Similarity 73.3%; Pred. No. 5.1e-55;
 Matches 486; Conservative 0; Mismatches 141; Indels 36; Gaps 10;

183 AGGATCTCCAGCTGTTACCTCCAGCCCGG-GCATTTCCAGGCGGGGATCTCCAGCC 241
 53 ATGATCTCCAGCAAGAACCTTTCAGCTGAGCATCTCCAGCCAGGCAATCTCCAGCT 112
 242 CAGCATCTCCAGCCCGGATCTCCGCTCTGGCATCTTCCAGGCTCTCATTC-AG 300
 113 GGGACCTCCAGGCGGGGATCTCCAGCCAGGATCATCTTCCAGGCTCTCATTC 172
 301 CAGGTATCATCTCCAGGCTCCAGGCTCCGCTGAGCAACCTCCCAACAGAGTATCTTGT 360
 173 CAGGTATCATCTCCAGGCTCCAGGCTCCGCTGAGCAACCTCCCAACAGAGTATCTTGT 232
 361 TAGAGCAACACAGTGGGGCTGTATCCCATCTCATCTCTCCAGAGTACAGCACAGC 420
 233 TAGAGCAACACAGTGGGGCTGTATCCCATCTCATCTCTCCAGAGTACAGCACAGC 292
 421 AACAGGGGCGCCAGGGGAGAGCCAGGATGAGAGCTGCGCAAGTTCACT--GGCGGGAG 478
 293 AACAGGGGCGCCAGGATGAGAGCCAGGATGAGAGCTGCGCAAGTTCACTGAGCAAGGAG 352
 479 GGGCAGAGAGAGCTACCGCTCATCGGCTGCGTGTCT--CCTCTCATTTGCTCTGTGG- 533
 353 GGGCAGAGAGAGCTACCGCTCATCGGCTGAGAGTGTCACTCTCATTTGCTCTGTGG- 412
 534 -TTTCGCTCATCTCTCTTCCAGTTCTGGAGAGGCG-----CACACAGGATCAGGTA 585
 413 TTTCGCTCATCTCTCTTCCAGTTCTGGAGAGGCGCACAGGATCAGGATCAGGTA 472
 586 CAAGAGAGAGAGAGCTGCTGCGCAAGCAG-CTGTTGCTGAGAGGGGTGGTGGACT 644
 473 GCAAGATGTGTGAAGCTTTGCTCCAAAGCAGCCTGTTGCTGAGAGGGGTGTATGG 532
 645 GCAAGATGTGTGAAGCTTTGCTCCAAAGCAGCCTGTTGCTGAGAGGGGTGTATGG 689
 533 GACTTCCAAAGATCTGAAGAGTGAAGAGCGCTGTGCTAGCGTATGACTGGAGC 592
 690 -ACAAGTCTCTGCTTAAATCTACTTGGGT-CTTCCATCAGTGGGCTTCCATCTGTAG 747
 593 AACCTCTCTGCTTAAATCTACTTGGGTCTTCCATCAGTGGGATCTGAGT 652
 748 CAGCAACT-GGAATGACTCTTCACTCAGAGAGAGCCTGCGAGAGTGGGTTTCGAGAGTG 806
 653 CAGCACTCGGGAATGACTCTTCACTCAGAGAGAGCCTGCGAGAGTGGGATCAAGAGGT 712
 807 CTC 809
 713 CCC 715

RESULT 14
 US-09-008-271A-18
 Sequence 18, Application US/0908271A
 Patent No. 6203979

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
 Hillman, Jennifer L.
 Yue, Henry
 Guegler, Karl J.
 Corley, Neil C.

Tang, Tom Y.
 Shah, Purvi
 TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/008,271A
 FILING DATE: 16-Jan-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: <Unknown>
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Mohan-Peterson, Sheela
 REGISTRATION NUMBER: 41,201
 REFERENCE/DOCKET NUMBER: PF-0458 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2038 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: COLN00113
 CLONE: 1337018
 SEQUENCE DESCRIPTION: SEQ ID NO: 18:
 US-09-008-271A-18

Query Match 11.1%; Score 193.2; DB 3; Length 2038;
 Best Local Similarity 55.8%; Pred. No. 3.1e-37;
 Matches 435; Conservative 0; Mismatches 333; Indels 12; Gaps 3;

924 TCTTCTCTCAGTGTCTCCCACTGCGGAGTGGGCAATGACCGGGGATCTGGGAGGG 983
 759 TCTTCTCTCAGTGTCTTGGCTGCGGAGGAGCTGGAAGACCCCGTGTGGTGGGG 818
 984 CGCTGCGCTGGAGTGAAGAGGCTTGGCAAGTGAAGTGTGCACTTGGGCAACCCACA 1043
 819 AGGAGGCTCTGTGAGATCTTGGCTTGGAGGATGACATTCAGTACGAACAAAGCAGC 878
 1044 TCTGTGAGGACAGCTCATTTGACGCGCAGTGGTGTCTCACTGCGGCCACTGCTTCTTG 1103
 879 TCTGTGAGGAGAGATCTTGAACCCCACTGGGTCTCAAGGAGCCCACTGCTTCTTCC 934
 1104 TGACCCGAGAGAGTCTTGAAGGCTGGAAGGTGTACGCGGAGCAACGACACTGCAAC 1163
 935 -AGGAACAATACCGATGTGTTCAACTGGAAGGTGCGGAGGCTCAGACAAACTGGGCA 992
 1164 AGTTCCTGAGAGGAGCTCATTTGGCGAGATCATCATCAACAGCAATTAACGATGAG 1223
 993 GCTTCC--ATCCTGTGTGTGGCAAGATCATCATTAATTAACCCCATGTACC 1049
 1224 AGGAGCACTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTGACCTGTCCGCTACA 1283
 1050 CCAAGACATATGACATGCGCTCATGAGAGCTGCAAGTCCCACTCTTCTCAGGACAG 1109
 1284 TCCACCTGTGCTGCTCCCATGATGACAGACTTTAGCTCAATGAGACCTGCTGGA 1343
 1110 TCAGGCCATCTGTCTCTCTTTGTATGAGAGGATCATCTCAGCAACCCCACTTGA 1169

| | | | |
|----|------|--|------|
| QY | 1344 | TCACAGGCTTTGGCAAGACCAAGGGAGACAGATGACAAGACATCCCCCTCTCCGGAGG | 1403 |
| Db | 1170 | TCATTGGATGGGGCTTTACGAAAGCAGATGAGGAGAGATGTCTGACATACTGCTCAGG | 1229 |
| QY | 1404 | TGCAGCTCAATCTCATCGACTTTCAGAAATGCATGACTACTTGGTCTATGACAGTTACC | 1463 |
| Db | 1230 | CGTACAGTTCAGGTCAATTGACAGCACACGGTCCATTGACAGACGATGCGTACCAAGGGAGG | 1289 |
| QY | 1464 | TTACCCCAAGAGATGATGTGCTGGGGACCTTGTGGGGGACAGACTCTGCCAGGGAG | 1523 |
| Db | 1290 | TCACCGAAGAGATGATGTGTGTCAGGCAATCCCGAAGGGGGTGTGGACACCTGCCAGGGTG | 1349 |
| QY | 1524 | ACAGCGGGGGGCCCTCTTGTCTGTGAGCAGAAACAACCGCTGTACTGGCAGGTGTACCA | 1583 |
| Db | 1350 | ACAGTGTGGGGCCCCCTGATGT---ACCAATCTGACACAGTGCATGTGTGGGCACTGTTA | 1406 |
| QY | 1584 | GCTGGGGCACAGGCTGTGGGCCAGAGAAACAACCTGTGTGTATACCAAAATGACAGAG | 1643 |
| Db | 1407 | GCTGGGGCTATGCGTGGGGGGGGCCCGAGCACCCAGAGATTAACACCAAGGTCTCACCT | 1466 |
| QY | 1644 | TTCTTCCCTGGATTTCACGAAAGATGGAGACGAGGTGCATTCAGAAAATCTCAACAG | 1703 |
| Db | 1467 | ATCTCACTGGAATCTACATATCTGGGAAGGCTGAGCTGTATGCTGCTGCCCTTTGGAG | 1526 |

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New). Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

| Result | Query | Match | Length | DB | ID | Description |
|--------|--------|-------|--------|----|--------------------|--------------------|
| ✓ No. | Score | | | | | |
| 1 | 1748 | 100.0 | 1748 | 3 | US-09-879-797-11 | Sequence 11, App1 |
| 2 | 1748 | 100.0 | 1748 | 9 | US-10-806-370-11 | Sequence 11, App1 |
| 3 | 1689 | 96.6 | 1707 | 8 | US-10-428-275-129 | Sequence 129, App |
| 4 | 1689 | 96.6 | 1708 | 8 | US-10-428-275-117 | Sequence 117, App |
| 5 | 1687.4 | 96.5 | 1689 | 7 | US-10-156-2184-23 | Sequence 23, App1 |
| 6 | 1687.4 | 96.5 | 1689 | 10 | US-10-916-1758-93 | Sequence 93, App1 |
| 7 | 1687.4 | 95.6 | 2033 | 7 | US-10-353-690-99 | Sequence 99, App1 |
| 8 | 1670.4 | 95.6 | 2067 | 7 | US-10-156-2184-25 | Sequence 25, App1 |
| 9 | 1663.4 | 95.2 | 2067 | 10 | US-10-916-1758-1 | Sequence 1, App11 |
| 10 | 1663.4 | 95.1 | 2432 | 8 | US-10-428-275-103 | Sequence 103, App |
| 11 | 1643.8 | 94.0 | 2432 | 8 | US-10-428-275-139 | Sequence 139, App |
| 12 | 1643.8 | 94.0 | 2432 | 8 | US-10-428-275-141 | Sequence 141, App |
| 13 | 1643.8 | 94.0 | 2432 | 8 | US-10-428-275-143 | Sequence 143, App |
| 14 | 1640 | 93.8 | 1771 | 8 | US-10-428-275-123 | Sequence 123, App |
| 15 | 1640 | 93.8 | 1779 | 8 | US-10-428-275-119 | Sequence 119, App |
| 16 | 1479 | 84.6 | 2310 | 8 | US-10-428-275-157 | Sequence 157, App |
| 17 | 1479 | 84.6 | 3324 | 10 | US-10-955-157-5559 | Sequence 2559, App |

| | | | | | | |
|----|--------|------|------|----|-------------------|-------------------|
| 18 | 1429 | 81.8 | 1614 | 3 | US-09-888-645-45 | Sequence 45, Appl |
| 19 | 1429 | 81.8 | 1614 | 15 | US-11-037-273-45 | Sequence 45, Appl |
| 20 | 1369 | 78.3 | 1626 | 8 | US-10-428-275-121 | Sequence 12, App |
| 21 | 1302.8 | 74.5 | 1314 | 3 | US-09-898-837A-17 | Sequence 14, Appl |
| 22 | 1302.8 | 74.5 | 1314 | 3 | US-09-898-837A-17 | Sequence 17, Appl |
| 23 | 1302.8 | 74.5 | 1314 | 8 | US-10-428-275-131 | Sequence 13, App |
| 24 | 1277.8 | 73.1 | 1574 | 7 | US-10-274-659-19 | Sequence 29, Appl |
| 25 | 1277.8 | 73.1 | 1574 | 8 | US-10-333-574-29 | Sequence 29, Appl |
| 26 | 1258.6 | 72.0 | 1434 | 6 | US-10-177-661-1 | Sequence 1, Appl |
| 27 | 1258.6 | 72.0 | 1434 | 10 | US-10-910-507-1 | Sequence 1, Appl |
| 28 | 1144.4 | 65.5 | 1146 | 8 | US-10-428-275-133 | Sequence 13, App |
| 29 | 1120.4 | 64.1 | 1218 | 8 | US-10-428-275-125 | Sequence 135, App |
| 30 | 1120.4 | 64.1 | 1218 | 8 | US-10-428-275-137 | Sequence 137, App |
| 31 | 1105 | 63.2 | 1203 | 8 | US-10-428-275-111 | Sequence 117, App |
| 32 | 1103.4 | 63.1 | 1203 | 8 | US-10-428-275-117 | Sequence 117, App |
| 33 | 1101.8 | 63.0 | 1203 | 8 | US-10-428-275-115 | Sequence 115, App |
| 34 | 1068.4 | 61.1 | 1078 | 3 | US-09-898-837A-16 | Sequence 16, Appl |
| 35 | 1062.6 | 60.8 | 1341 | 6 | US-10-177-661-3 | Sequence 3, Appl |
| 36 | 1062.6 | 60.8 | 1341 | 10 | US-10-910-507-3 | Sequence 3, Appl |
| 37 | 1053 | 60.2 | 1230 | 3 | US-09-879-792-35 | Sequence 35, Appl |
| 38 | 1053 | 60.2 | 1230 | 9 | US-10-806-370-35 | Sequence 35, Appl |
| 39 | 909 | 52.0 | 1110 | 8 | US-10-428-275-113 | Sequence 113, App |
| 40 | 906.2 | 51.8 | 1101 | 8 | US-10-428-275-109 | Sequence 109, App |
| 41 | 782.4 | 44.8 | 1222 | 3 | US-09-804-156-5 | Sequence 5, Appl |
| 42 | 782.4 | 44.8 | 1222 | 3 | US-09-946-633-3 | Sequence 3, Appl |
| 43 | 782.4 | 44.8 | 1222 | 6 | US-10-125-459-3 | Sequence 3, Appl |
| 44 | 782.4 | 44.8 | 1222 | 6 | US-10-067-761-5 | Sequence 5, Appl |
| 45 | 782.4 | 44.8 | 1222 | 7 | US-10-315-519-5 | Sequence 5, Appl |

ALIGNMENTS

```

RESULT 1
US-09-879-792-11
: Sequence 11, Application US/098879792
: Patent No. US20020061850A1
: GENERAL INFORMATION:
: APPLICANT: Xiao, Yonghong
: APPLICANT: Gedrich, Richard
: TITLE OF INVENTION: Regulation of Human Transmembrane Serine
: TITLE OF INVENTION: Protease
: FILE REFERENCE: 02973.00035
: CURRENT APPLICATION NUMBER: US/09/879,792
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/211,224
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: US 60/283,353
: PRIOR FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/283,648
: PRIOR FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: PCT _____
: PRIOR FILING DATE: 2001-06-12
: (Docket No. US20020061850A1 LIO-81-WO)
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 1748
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-879-792-11

```

| | | | | | | |
|----|-----------------------|---|---------------|-----------|--------------|--|
| | Query Match | 100.0% | Score 1748; | DB 3; | Length 1748; | |
| | Best Local Similarity | 100.0% | Pred. No. 0; | | | |
| | Matches 1748; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; | |
| Oy | 1 | CTCAGAGCATCGAGGAGGCACGCCAATGCATCTCCAGCAAGAACCCTTCAGC | 60 | | | |
| | | | | | | |
| Dd | 1 | CTCGAGGACCTTGAGAGGGGACAGCCACGGGAATGCATCTCCAGCAAGAACCTTCAGC | 60 | | | |
| Oy | 61 | TGAGAGCATCTCCAGCCGAGGCATCTCCAGCTGGGACA C CTCAGGCCGGGATCTCCAGC | 120 | | | |
| | | | | | | |
| Dd | 61 | TGAGAGCATCTCCAGCCGAGGCATCTCCAGCTGGGACA C CTCAGGCCGGGATCTCCAGC | 120 | | | |

```

Oy 121 CCAGGATCTCCAGGCCAGGATCTCCAGTGGAGACACTCCGGGCGGAGATCTCCAGC 180
Db 121 CCAGGATCTCCAGGCCAGGATCTCCAGTGGAGACACTCCGGGCGGAGATCTCCAGC 180
Oy 181 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 240
Db 181 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 240
Oy 241 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 300
Db 241 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 300
Oy 301 CAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 360
Db 301 CAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 360
Oy 361 TAGAGCAACACAGTGGGGGCTGTACCCATCCGATCATCTCCGCAAGTGAAGTGAAGTGA 420
Db 361 TAGAGCAACACAGTGGGGGCTGTACCCATCCGATCATCTCCGCAAGTGAAGTGAAGTGA 420
Oy 421 AACAGGAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 421 AACAGGAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Oy 481 CAGAGAGAGTACCGCTCATCGAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 481 CAGAGAGAGTACCGCTCATCGAGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Oy 541 CATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 541 CATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Oy 601 GAGCTGTCCAGAGCAAGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 601 GAGCTGTCCAGAGCAAGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Oy 661 CAGCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 661 CAGCTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Oy 721 CTCCATCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 721 CTCCATCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Oy 781 CTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 781 CTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Oy 841 TGCCAAAGCTTCTCAATCTTGAATACCACTCCCACTCCCACTCCCACTCCCACTCCCACT 900
Db 841 TGCCAAAGCTTCTCAATCTTGAATACCACTCCCACTCCCACTCCCACTCCCACTCCCACT 900
Oy 901 TGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 901 TGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Oy 961 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 961 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Oy 1021 TCTGCACTTCCGAGCAACCTCATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1021 TCTGCACTTCCGAGCAACCTCATCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Oy 1081 CACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1081 CACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Oy 1141 GCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1141 GCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Oy 1201 CAACGCAATTAACCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260

```

```

Db 1201 CAACGCAATTAACCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Oy 1261 GCGCGTACCTGTCCGATCATCTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1261 GCGCGTACCTGTCCGATCATCTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Oy 1321 TAGGCTCAATGAGACTGTGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Db 1321 TAGGCTCAATGAGACTGTGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Oy 1381 GACATCCGCTTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1381 GACATCCGCTTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Oy 1441 CTACTGCTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 1441 CTACTGCTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Oy 1501 GGGAGAGAGCTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db 1501 GGGAGAGAGCTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Oy 1561 CTGCTACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Db 1561 CTGCTACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Oy 1621 TGATGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Db 1621 TGATGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Oy 1681 GCGATTCAGAAATCTTAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1681 GCGATTCAGAAATCTTAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Oy 1741 TCGAGAAA 1748
Db 1741 TCGAGAAA 1748

```

```

RESULT 2
US-10-806-370-11
; Sequence 11, Application US/10806370
; Publication No. US20040209327A1
; GENERAL INFORMATION:
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/10/806.370
; CURRENT FILING DATE: 2004-03-23
; PRIOR APPLICATION NUMBER: US/09/879.792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211.224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283.353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283.648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1748
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-806-370-11
Query Match 100.0%; Score 1748; DB 9; Length 1748;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


QY 1 CTGAGAGACATGAGAGAGGAGCAGGCAAGGGAATGCACTTCAGCAAGAACACCTTGAGC 60
DB 1 CTGAGAGACATGAGAGAGGAGCAGGCAAGGGAATGCACTTCAGCAAGAACACCTTGAGC 60
QY 61 TGGAGCATCTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 120
DB 61 TGGAGCATCTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 120
QY 121 CCAGGCACTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 180
DB 121 CCAGGCACTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 180
QY 181 CCAGGCACTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 240
DB 181 CCAGGCACTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 240
QY 241 CCAGGCACTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 300
DB 241 CCAGGCACTTCAGAGCCAGGATCTCCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGC 300
QY 301 CAGGTCATCATCCGCGAGTCAGGCTCGGTCAGCAACTTCCCAACAGAGTGAACCTTGT 360
DB 301 CAGGTCATCATCCGCGAGTCAGGCTCGGTCAGCAACTTCCCAACAGAGTGAACCTTGT 360
QY 361 TAGAGCAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 361 TAGAGCAACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 421 AACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 421 AACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 481 CCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 481 CCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 541 CATATCTCTCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 CATATCTCTCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 601 GAGCTGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 601 GAGCTGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 661 CGAGCTGGGCTGCGTAGAGTTGACTGGGACAAAGTCTCTGCTTAAATCTACTCTGGGTC 720
DB 661 CGAGCTGGGCTGCGTAGAGTTGACTGGGACAAAGTCTCTGCTTAAATCTACTCTGGGTC 720
QY 721 CTCCATCATGAGGCTTCCATCTGTAGAGCACTGGAATGATCTTACTGAGAGAGAC 780
DB 721 CTCCATCATGAGGCTTCCATCTGTAGAGCACTGGAATGATCTTACTGAGAGAGAC 780
QY 781 CTGCGAGAGGCTGGGATTCGAGAGTCTCAACGAGCAACGAGGTTGGCCACAGGAGATT 840
DB 781 CTGCGAGAGGCTGGGATTCGAGAGTCTCAACGAGCAACGAGGTTGGCCACAGGAGATT 840
QY 841 TGCCAGAGCTTCTCAATCTTGAGATCACTCAACATCCAGAGAAAGCTTCCAGAGTC 900
DB 841 TGCCAGAGCTTCTCAATCTTGAGATCACTCAACATCCAGAGAAAGCTTCCAGAGTC 900
QY 901 TGAATGCCCTTCCAGAGGATTCCTCCAGAGTCTCCAGAGTCTCCAGAGTCTCCAGAG 960
DB 901 TGAATGCCCTTCCAGAGGATTCCTCCAGAGTCTCCAGAGTCTCCAGAGTCTCCAGAG 960
QY 961 GAGCGGAGGATCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 961 GAGCGGAGGATCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1021 TGTGCACTTGGGAGCAACCAATCTGTGAGAGCAGCTCAATGACGCCAGTGGTGCT 1080
DB 1021 TGTGCACTTGGGAGCAACCAATCTGTGAGAGCAGCTCAATGACGCCAGTGGTGCT 1080

QY 1081 CACTGCCGCCACTGCTCTTCTTGAGACCCGAGAGAAAGTCTTGAGAGGCTGGAAGTGTGA 1140
DB 1081 CACTGCCGCCACTGCTCTTCTTGAGACCCGAGAGAAAGTCTTGAGAGGCTGGAAGTGTGA 1140
QY 1141 CGGAGGAGCAAGCACTGCAACGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1141 CGGAGGAGCAAGCACTGCAACGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1201 CAACAGCAATTTACACGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1201 CAACAGCAATTTACACGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1261 GCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
DB 1261 GCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
QY 1321 TAGCTCAATGAGACCTGCTGATCAGAGGCTTTGGCAAGACAGGAGAGCAGATGACAA 1380
DB 1321 TAGCTCAATGAGACCTGCTGATCAGAGGCTTTGGCAAGACAGGAGAGCAGATGACAA 1380
QY 1381 GACATCCCCCTCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1381 GACATCCCCCTCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1441 CTACTTGTATGACAGTTAATCTTACCCCAAGATGATGCTGAGGAGGAGGAGGAGGAGGAGGAGG 1500
DB 1441 CTACTTGTATGACAGTTAATCTTACCCCAAGATGATGCTGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 1501 GGGGAGAGACTCTCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
DB 1501 GGGGAGAGACTCTCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
QY 1561 CTGATACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 1561 CTGATACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
QY 1621 TGTGTACACCAAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
DB 1621 TGTGTACACCAAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
QY 1681 GCGATTGAGAAATCTTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
DB 1681 GCGATTGAGAAATCTTACCAAGG 1740
QY 1741 TCGAGAAA 1748
DB 1741 TCGAGAAA 1748

RESULT 3
US-10-428-275-129
; Sequence 129, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; PRIOR FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Curaseq1ist version 0.1
SEQ ID NO 129
LENGTH: 1707
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ..(1707)
US-10-428-275-129

Query Match 96.6%; Score 1689; DB 8; Length 1707;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ACCATTGAGAGGACAGCCAGGAAATGATCTCCAGCAGAAACACCTTCAGCTGAGCA 67
DB 10 ACCATTGAGAGGAGCAGCCAGGAAATGATCTCCAGCAGAAACACCTTCAGCTGAGCA 69
QY 68 TCTCCAGCCGAGGATCTCCAGCTGGGACACTCCAGCCGAGGATCTCCAGCCAGGCA 127
DB 70 TCTCCAGCCGAGGATCTCCAGCTGGGACACTCCAGCCGAGGATCTCCAGCCAGGCA 129
QY 128 TCTCCAGCCGAGGATCTCCAGCTGGGACACTCCAGCCGAGGATCTCCAGCCAGGCA 187
DB 130 TCTCCAGCCGAGGATCTCCAGCTGGGACACTCCAGCCGAGGATCTCCAGCCAGGCA 189
QY 188 TCTCCAGCTGTATACCTCCAGCCGAGGATCTCCAGCCGAGGATCTCCAGCCAGGCA 247
DB 190 TCTCCAGCTGTATACCTCCAGCCGAGGATCTCCAGCCGAGGATCTCCAGCCAGGCA 249
QY 248 TCTCCAGCCGAGGATCTCCAGCTGGGACACTCCAGCTGTATACCTCCAGCTGTATAC 307
DB 250 TCTCCAGCCGAGGATCTCCAGCTGGGACACTCCAGCTGTATACCTCCAGCTGTATAC 309
QY 308 TCAATCCGAGGATCTCCAGCTGGGACACTCCAGCTGTATACCTCCAGCTGTATAC 367
DB 310 TCAATCCGAGGATCTCCAGCTGGGACACTCCAGCTGTATACCTCCAGCTGTATAC 369
QY 368 ACACCAATGGGGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 427
DB 370 ACACCAATGGGGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 429
QY 428 GCCACAGGAGAGCCAGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 487
DB 430 GCCACAGGAGAGCCAGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 489
QY 488 CAGCTACCGCTCATCGGGTGGGCTCTCTCTCAATGGCTGTATGCTGTATGCTCATCATC 547
DB 490 CAGCTACCGCTCATCGGGTGGGCTCTCTCTCAATGGCTGTATGCTGTATGCTCATCATC 549
QY 548 CTCTTCCAGTTTGGAGAGGCAACAGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 607
DB 550 CTCTTCCAGTTTGGAGAGGCAACAGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 609
QY 608 CCCAAGCAGCTGTCTCTGTGACGAGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 667
DB 610 CCCAAGCAGCTGTCTCTGTGACGAGGATGTAACCTCCAGCTGTATACCTCCAGCTGTATACCTCCAGCTGTATAC 669
QY 668 GAGTGCCTGAGGTTTGAAGTGGGACAGCTCTGCTTAATCTAATCTGAGTCTCTCCAT 727
DB 670 GAGTGCCTGAGGTTTGAAGTGGGACAGCTCTGCTTAATCTAATCTGAGTCTCTCCAT 729
QY 728 CAGTGCCTTCCATCTGTAGAGCACTGGAATGATCTCTAAGAGAAAGCTTCCAG 787
DB 730 CAGTGCCTTCCATCTGTAGAGCACTGGAATGATCTCTAAGAGAAAGCTTCCAG 789
QY 788 CAGCTGGGTTTGGAGAGTGTACCGGACAGAGGTTTCCAGGAGTTTGGCAAC 847
DB 790 CAGCTGGGTTTGGAGAGTGTACCGGACAGAGGTTTCCAGGAGTTTGGCAAC 849

QY 848 AGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCTTCAAGTCTGAATGC 907
DB 850 AGCTTCTCAATCTTGAGATACAACTCCAGCATCCAGGAAAGCTTCAAGTCTGAATGC 909
QY 908 CTTTCCAGGAGGATATCTCCCTCAAGTGTTCCTCACTGCGGATCTGAGGAGGCAATGACCGGG 967
DB 910 CTTTCCAGGAGGATATCTCCCTCAAGTGTTCCTCACTGCGGATCTGAGGAGGCAATGACCGGG 969
QY 968 CGGATCGTGGAGGAGGAGCTGAGCTTCCGATATGCAAGTGGCTTGGCAAGTGTCTGCAC 1027
DB 970 CGGATCGTGGAGGAGGAGGAGCTGAGCTTCCGATATGCAAGTGGCTTGGCAAGTGTCTGCAC 1029
QY 1028 TTGGGACCAACCCACATCTGTGAGGACGCTCATTTAGCCGAGTGGTCTCACTGCC 1087
DB 1030 TTGGGACCAACCCACATCTGTGAGGACGCTCATTTAGCCGAGTGGTCTCACTGCC 1089
QY 1088 GCCCACTGCTTCTTCTGAGACCCGAGGAAAGTCTGTGAGGCTTGGAGGCTGTAGCGGGC 1147
DB 1090 GCCCACTGCTTCTTCTGAGACCCGAGGAAAGTCTGTGAGGCTGTAGCGGGC 1149
QY 1148 ACCAGCAACCTGACACAGTTGCTGAGGACGCTCATTTGCGGATCATCATCAAGC 1207
DB 1150 ACCAGCAACCTGACACAGTTGCTGAGGACGCTCATTTGCGGATCATCATCAAGC 1209
QY 1208 AATTACACCGATGAGAGAGCACTATGATAGCTCTCATGCGGCTGTCCAGCCCTG 1267
DB 1210 AATTACACCGATGAGAGAGCACTATGATAGCTCTCATGCGGCTGTCCAGCCCTG 1269
QY 1268 ACCCTGTCGCTCATATCCACCCGCTTGGCTCCCAATGAGATGAGACCTTTAGCCTC 1327
DB 1270 ACCCTGTCGCTCATATCCACCCGCTTGGCTCCCAATGAGATGAGACCTTTAGCCTC 1329
QY 1328 AATGAGACTGCTGATACAGGCTTTGGCAAGACAGGAGAGACAGATGACAAAGATCC 1387
DB 1330 AATGAGACTGCTGATACAGGCTTTGGCAAGACAGGAGAGACAGATGACAAAGATCC 1389
QY 1388 CCTTCTCCGAGAGGTGACAGTCAATCTATGATGACTTCAAGAAATGATGACTTGG 1447
DB 1390 CCTTCTCCGAGAGGTGACAGTCAATCTATGATGACTTCAAGAAATGATGACTTGG 1449
QY 1448 GTCTATGACAGTTACCTTACCCCAAGATGATGCTGGGAGCTTGTGGGGGAGAG 1507
DB 1450 GTCTATGACAGTTACCTTACCCCAAGATGATGCTGGGAGCTTGTGGGGGAGAG 1509
QY 1508 GACTCTGCAAGGAGACAGCGGGGAGCTTGTGTGTGAGCAGAAACACCGTGTGTAC 1567
DB 1510 GACTCTGCAAGGAGACAGCGGGGAGCTTGTGTGTGAGCAGAAACACCGTGTGTAC 1569
QY 1568 CTGGCAGGTGTACACAGCTGGGACACAGCTGTGTGAGCAGAAACACCTGTGTGTAC 1627
DB 1570 CTGGCAGGTGTACACAGCTGGGACACAGCTGTGTGAGCAGAAACACCTGTGTGTAC 1629
QY 1628 ACCAAGTGAAGAGTTCTTCCCTGATTTTACGAAATGAGAGCGAGGTGCAATTG 1687
DB 1630 ACCAAGTGAAGAGTTCTTCCCTGATTTTACGAAATGAGAGCGAGGTGCAATTG 1689
QY 1688 AGAAATCC 1696
DB 1690 AGAAATCC 1698

RESULT 4
US-10-428-275-147/c
Sequence 147, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD;
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428, 275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Curaseqlist version 0.1
SEQ ID NO 147
LENGTH: 1708
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (849)
US-10-428-275-147

Query Match 96.6%; Score 1689; DB 8; Length 1708;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

8 ACCATGAGAGAGGACAGCCACGGGAAATCATCTCCAGCAAGAACCTTCAGCTGAGACA 67
1698 ACCATGAGAGAGGACAGCCACGGGAAATCATCTCCAGCAAGAACCTTCAGCTGAGACA 1639
68 TCTCCAGCCCGGAGCATCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCA 127
1638 TCTCCAGCCCGGAGCATCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCA 1579
128 TCTCCAGCCCGGAGCATCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCA 187
1578 TCTCCAGCCCGGAGCATCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCA 1519
188 TCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCGGGACATCTCCAGGCCA 247
1518 TCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCGGGACATCTCCAGGCCA 1459
248 TCTCCAGCCCGGAGCATCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCA 307
1458 TCTCCAGCCCGGAGCATCTCCAGCTGAGACACCTCCAGGCCGGGACATCTCCAGGCCA 1399
308 TCAATCCGCGAGGTGAGCTCCGAGTGAACAACCTCCCAACAGAGTGAACCTTGTAGACA 367
1398 TCAATCCGCGAGGTGAGCTCCGAGTGAACAACCTCCCAACAGAGTGAACCTTGTAGACA 1339
368 ACACGAGTGGGGGCTGTAACCATCCGATCATCTCTGCAAGTGAAGCAACGACGAG 427
1338 ACACGAGTGGGGGCTGTAACCATCCGATCATCTCTGCAAGTGAAGCAACGACGAG 1279
428 GCCACGAGGAGAGCCACGAGTGAAGAGCTGACCAAGTTCACCTGGCGGGAGGCCAGAG 487
1278 GCCACGAGGAGAGCCACGAGTGAAGAGCTGACCAAGTTCACCTGGCGGGAGGCCAGAG 1219
488 CAGCTACCGCTCATCTGAGGTGCTGCTCTCTCATTTGCTGCTGCTTCTCATCTC 547
1218 CAGCTACCGCTCATCTGAGGTGCTGCTCTCTCATTTGCTGCTGCTTCTCATCTC 1159
548 CTCTTCCAGTCTGAGCAGGACCAACAGAGTCAAGTCAAGAGCAGAGGAGAGCTGT 607
1158 CTCTTCCAGTCTGAGCAGGACCAACAGAGTCAAGTCAAGAGCAGAGGAGAGCTGT 1099
608 CCAAGACGCTGCTGCTGAGCGGGTGTGAGCTGCAAGCTGAAGAGTGAAGAGCTG 667
1098 CCAAGACGCTGCTGCTGAGCGGGTGTGAGCTGCAAGCTGAAGAGTGAAGAGCTG 1039

668 GGCTGCGTGAAGTTTGAAGTGGAGCAAGTCTCTGTTAAATCTACTCTGGTCTCTCCAT 727
1038 GGCTGCGTGAAGTTTGAAGTGGAGCAAGTCTCTGTTAAATCTACTCTGGTCTCTCCAT 979
728 CAGTGGCTTCCCATCTGTAAGAGCAACTGGAATGACTCTTCACTAGAGAAAGCTTCCAG 787
978 CAGTGGCTTCCCATCTGTAAGAGCAACTGGAATGACTCTTCACTAGAGAAAGCTTCCAG 919
788 CAGTGGCTTCCCATCTGTAAGAGCAACTGGAATGACTCTTCACTAGAGAAAGCTTCCAG 847
918 CAGTGGCTTCCCATCTGTAAGAGCAACTGGAATGACTCTTCACTAGAGAAAGCTTCCAG 859
848 AGCTTCTCAATCTTGAAGATCAACTCCACCATCCAGAGAAAGCTTCCAGAGTCTGAATGC 907
858 AGCTTCTCAATCTTGAAGATCAACTCCACCATCCAGAGAAAGCTTCCAGAGTCTGAATGC 799
908 CTTTCCAGCGGTATATCTCCCTCCAGTGTCCCACTGCGGACTGAGGCGCATGACCGGG 967
798 CTTTCCAGCGGTATATCTCCCTCCAGTGTCCCACTGCGGACTGAGGCGCATGACCGGG 739
968 CGATTCGTGGAGGGGGGCTGCTCGGATGAGCAAGTGGCTTGGCAAGTGAAGTCTGAC 1027
738 CGATTCGTGGAGGGGGGCTGCTCGGATGAGCAAGTGGCTTGGCAAGTGAAGTCTGAC 679
1028 TTGGGACCAACCAATCTGTAAGAGCAAGCTCATTTGACGCGCAAGTGGTCTCACTGCC 1087
678 TTGGGACCAACCAATCTGTAAGAGCAAGCTCATTTGACGCGCAAGTGGTCTCACTGCC 619
1088 GCCCACTGCTTCTTCTGTAAGCGGGAGAGGCTCTGAGAGGCTGGAAGTGAAGCGGGC 1147
618 GCCCACTGCTTCTTCTGTAAGCGGGAGAGGCTCTGAGAGGCTGGAAGTGAAGCGGGC 559
1148 ACCGAACTCTGACCAAGTGTCTGAGAGAGCTCTCACTTGGGAGTATATCAACAGC 1207
558 ACCGAACTCTGACCAAGTGTCTGAGAGAGCTCTCACTTGGGAGTATATCAACAGC 499
1208 AATTACACCATGAGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCAGAGCCCTG 1267
498 AATTACACCATGAGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCAGAGCCCTG 439
1268 ACCCTGTCGCTCAATCCACCTGCTTGCCTCCCATGATGAGACAGACCTTTAGCCTC 1327
438 ACCCTGTCGCTCAATCCACCTGCTTGCCTCCCATGATGAGACAGACCTTTAGCCTC 379
1328 AATGAGACTGCTGATCAAGGCTTGGCAAGACAGAGAGACAGATGACAGACATCC 1387
378 AATGAGACTGCTGATCAAGGCTTGGCAAGACAGAGAGACAGATGACAGACATCC 319
1388 CCCTTCCCGGGAGGAGTGAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTGG 1447
318 CCCTTCCCGGGAGGAGTGAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTGG 259
1448 GTCTATGACAGTTACTTTACCCCAAGATGATGTGTGCTGGGAGCTTCTGTTGGGGCAGA 1507
258 GTCTATGACAGTTACTTTACCCCAAGATGATGTGTGCTGGGAGCTTCTGTTGGGGCAGA 199
1508 GACTCTGCGAGGAGACAGCGGGGGCTTGTCTGTGAGCAAGAACACCGCTGTAAC 1567
198 GACTCTGCGAGGAGACAGCGGGGGCTTGTCTGTGAGCAAGAACACCGCTGTAAC 139
1568 CTGGCAGGTGTCACAGCTGGGGGACAGGCTGTGGGCAAGAAACAACTGGTGTGAC 1627
138 CTGGCAGGTGTCACAGCTGGGGGACAGGCTGTGGGCAAGAAACAACTGGTGTGAC 79
1628 ACCAAGTGAAGAAATCTTCTCTGATTTTACAGCAAGATGAGAGAGAGTGCATTC 1687
78 ACCAAGTGAAGAAATCTTCTCTGATTTTACAGCAAGATGAGAGAGAGTGCATTC 19
1688 AGAAATCC 1696
18 AGAAATCC 10

RESULT 5
US-10-156-214A-23
; Sequence 23, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempie
; APPLICANT: George P. Vlasuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Siew
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; TITLE OF INVENTION: Theoret
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156, 214A
; CURRENT FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)...(1689)
; OTHER INFORMATION: Nucleic acid encoding Endothelinase 2-S protein
US-10-156-214A-23

Query Match Best Local Similarity 96.5%; Score 1687.4; DB 7; Length 1689;

Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 ATGGAGAGGACACGACCGGGAATGATCTTCACGAAACACCTTCAGCTGAGCATCT 70
Db 1 ATGGAGAGGACACGACCGGGAATGATCTTCACGAAACACCTTCAGCTGAGCATCT 60
QY 71 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 130
Db 61 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 120
QY 131 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 190
Db 121 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 180
QY 191 CCAGCTGTACACCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGCCAGCATCT 250
Db 181 CCAGCTGTACACCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGCCAGCATCT 240
QY 251 CCAGCCCGGGATCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGCCAGCATCT 310
Db 241 CCAGCCCGGGATCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGCCAGCATCT 300
QY 311 TCAGCCAGGATCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGAGCCAGCATCT 370
Db 301 TCAGCCAGGATCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGAGCCAGCATCT 360
QY 371 CCAGTGGAGGCTGTACCATTCATCTTCCTCCAGGTTCAGGATTCAGGACCAAGGAGCC 430
Db 361 CCAGTGGAGGCTGTACCATTCATCTTCCTCCAGGTTCAGGATTCAGGACCAAGGAGCC 420
QY 431 ACCAGGAGAGGACCAAGGATTCAGAGCCGGGATTCAGAGCCGGGATTCAGAGGAGCATCT 490
Db 421 ACCAGGAGAGGACCAAGGATTCAGAGCCGGGATTCAGAGCCGGGATTCAGAGGAGCATCT 480
QY 491 CTACCGCTCATCGGAGTGGGCTCTCTCTCATTTGCGCTGTGATTCATCTCTCT 550
Db 481 CTACCGCTCATCGGAGTGGGCTCTCTCTCATTTGCGCTGTGATTCATCTCTCT 540
QY 551 TTCCAGTTCTTGAGAGGACCAAGGATTCAGAGCCGGGATTCAGAGGAGGAGGAGCTGTCC 610
Db 541 TTCCAGTTCTTGAGAGGACCAAGGATTCAGAGCCGGGATTCAGAGGAGGAGGAGCTGTCC 600
QY 611 AAGCAGCGCTGTCCCTGTGACGGGAGTGGATTCGCAAGCTGAAGAGTGAAGCTGGCC 670

Db 601 AAGCAGCGCTGTCCCTGTGACGGGAGTGGATTCGCAAGCTGAAGAGTGAAGCTGGCC 660
QY 671 TGCGTGAAGTTTGAAGGACCAAGTCTCTGTTAAATCTACTCTGGGCTCTCCATCAG 730
Db 661 TGCGTGAAGTTTGAAGGACCAAGTCTCTGTTAAATCTACTCTGGGCTCTCCATCAG 720
QY 731 TGCGTTCATCTGTAGACCAAGTGAATGATCTCTACTCAGAGAAAGCTGCGACAG 790
Db 721 TGCGTTCATCTGTAGACCAAGTGAATGATCTCTACTCAGAGAAAGCTGCGACAG 780
QY 791 CTGGGTTTCCAGAGTGTCTACCCGGAACCAAGGTTTCCCAAGAGGATTTTCCAAAGC 850
Db 781 CTGGGTTTCCAGAGTGTCTACCCGGAACCAAGGTTTCCCAAGAGGATTTTCCAAAGC 840
QY 851 TTCTCAATCTTGAGATACAACTCCACCATCCAGAAAGCTCCACAGGTGAAATGCGCT 910
Db 841 TTCTCAATCTTGAGATACAACTCCACCATCCAGAAAGCTCCACAGGTGAAATGCGCT 900
QY 911 TCCAGCGGATATCTCCCTCCAGTGTCCCACTGCGGACTGAGAGGCTATGACCGGCGG 970
Db 901 TCCAGCGGATATCTCCCTCCAGTGTCCCACTGCGGACTGAGAGGCTATGACCGGCGG 960
QY 971 ATCGTGGAGAGGAGGCTGGCTGGGATGCAAGGAGCTTGGCAAGTGAAGTGCATCTC 1030
Db 961 ATCGTGGAGAGGAGGCTGGCTGGGATGCAAGGAGCTTGGCAAGTGAAGTGCATCTC 1020
QY 1031 GGCAACCAACATCTGTGAGAGGACGCTCATTTGACCCGCAAGTGGTGTCTCATCGCGCC 1090
Db 1021 GGCAACCAACATCTGTGAGAGGACGCTCATTTGACCCGCAAGTGGTGTCTCATCGCGCC 1080
QY 1091 CACTGCTTTCTGTGACCCGAGAGAGGCTCTGAGAGGCTTGGAAAGTGTACCGGAGCAC 1150
Db 1081 CACTGCTTTCTGTGACCCGAGAGAGGCTCTGAGAGGCTTGGAAAGTGTACCGGAGCAC 1140
QY 1151 AGCAACCTTGACCAAGTTGCTGAGGAGGCTCTCATTTCCGAGATCATCAAGCAAT 1210
Db 1141 AGCAACCTTGACCAAGTTGCTGAGGAGGCTCTCATTTCCGAGATCATCAAGCAAT 1200
QY 1211 TACACGATGAGAGAGGACGATATGATGATGAGGCTCTCATGCGGCTTCAAGCCCTGACC 1270
Db 1201 TACACGATGAGAGAGGACGATATGATGATGAGGCTCTCATGCGGCTTCAAGCCCTGACC 1260
QY 1271 CTGTGCTCATCATCACCTCTGCTTGCCTCCCATGATGAGACAGACCTTTAGCTCAAT 1330
Db 1261 CTGTGCTCATCATCACCTCTGCTTGCCTCCCATGATGAGACAGACCTTTAGCTCAAT 1320
QY 1331 GAGACCTGTGATTCAGAGCTTTGGCAAGCAAGGAGAGCAATGACAAATCTCCGCC 1390
Db 1321 GAGACCTGTGATTCAGAGCTTTGGCAAGCAAGGAGAGCAATGACAAATCTCCGCC 1380
QY 1391 TTCTCCGAGAGGAGGACGATCAATCTCATGACCTTCAAGAAATGCAATGACTTGGTTC 1450
Db 1381 TTCTCCGAGAGGAGGACGATCAATCTCATGACCTTCAAGAAATGCAATGACTTGGTTC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGCGGACCTTCTGGGAGCAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGATGATGTGTGCGGACCTTCTGGGAGCAGAGAC 1500
QY 1511 TCTGCGAGGAGAGACAGCGGGGCTCTTGTCTGTAGAGCAACACCGTGTATCCTG 1570
Db 1501 TCTGCGAGGAGAGACAGCGGGGCTCTTGTCTGTAGAGCAACACCGTGTATCCTG 1560
QY 1571 GCAAGTGTCAACAGCTGGGACAGGCTGTGTGCGAGAGAAACCTGTGTGTACAC 1630
Db 1561 GCAAGTGTCAACAGCTGGGACAGGCTGTGTGCGAGAGAAACCTGTGTGTACAC 1620
QY 1631 AAAGTGAAGAAGTTCTTCCCTGATTTTACAGCAAGTGAAGAGGAGTGCATTGAG 1690
Db 1621 AAAGTGAAGAAGTTCTTCCCTGATTTTACAGCAAGTGAAGAGGAGTGCATTGAG 1680
QY 1691 AAATCTTAA 1699
|||||

Db 1681 AAATCTTA 1689

RESULT 6
US-10-916-758-93
; Sequence 93, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELIAL-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; PRIOR FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-916-758-93

Query Match 96.5%; Score 1687.4; DB 10; Length 1689;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

11 ATGAGAGGAGGACAGCCAGGGAATGATCTCTCAGCAAGAACACCTTCAGCTGAGCATCT 70
1 ATGAGAGGAGGACAGCCAGGGAATGATCTCTCAGCAAGAACACCTTCAGCTGAGCATCT 60
71 CCAGCCAGGATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 130
61 CCAGCCAGGATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 120
131 CCAGCCAGGATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 190
121 CCAGCCAGGATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 180
191 CCAGCTGTACATCTCCAGAGCCGGGATCTCCAGAGCCGGGATCTCCAGCCAGGATCT 250
181 CCAGCTGTACATCTCCAGAGCCGGGATCTCCAGAGCCGGGATCTCCAGCCAGGATCT 240
251 CCAGCCAGGATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 310
241 CCAGCCAGGATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 300
311 TCCGCGAGGTAGAGCTCGGTACAACTTCCCAACCAAGTGTACTTTTGAAGCAACA 370
301 TCCGCGAGGTAGAGCTCGGTACAACTTCCCAACCAAGTGTACTTTTGAAGCAACA 360
371 CCAGTGGGGGCTGATACCATCCGATCATCTCTGCGAGGTAGAGCAAGGAGGAGG 430
361 CCAGTGGGGGCTGATACCATCCGATCATCTCTGCGAGGTAGAGCAAGGAGGAGG 420
431 ACCAGGAGAGCCAGGATAGAGCTCGGCAAGTTCACTTGGCGGAGGAGGAGGAGG 490
421 ACCAGGAGAGCCAGGATAGAGCTCGGCAAGTTCACTTGGCGGAGGAGGAGGAGG 480
491 CTACCGCTCATCGGGTGTGCTCTCTCTCATTTGCTGCTGCTGCTCATCTCTC 550
481 CTACCGCTCATCGGGTGTGCTCTCTCTCATTTGCTGCTGCTGCTCATCTCTC 540
551 TTCCAGTTCTGGAGGAGGACACAGAGGATCAGGTACAAGAGAGAGAGGAGGAGG 610
541 TTCCAGTTCTGGAGGAGGACACAGAGGATCAGGTACAAGAGAGAGAGGAGGAGG 600
611 AAGCAGCGTGTCTGCTGTGACGGGGTGTGAGCTGCAAGCTGAAGAGTGAAGAGTGG 670
601 AAGCAGCGTGTCTGCTGTGACGGGGTGTGAGCTGCAAGAGTGAAGAGTGGAG 660

Qy 671 TGGTGAGGTTTGAAGTGGAGCAAGTCTCTGTTAAATTTACTGTGGGTCTCCCATCAG 730
Db 661 TGGTGAGGTTTGAAGTGGAGCAAGTCTCTGTTAAATTTACTGTGGGTCTCCCATCAG 720
Qy 731 TGGTTCCTCATCTGTAGACCAACTGGATGACTCTTACTGAGAGAGACTGCGACAG 790
Db 721 TGGTTCCTCATCTGTAGACCAACTGGATGACTCTTACTGAGAGAGACTGCGACAG 780
Qy 791 CTGGGTTTCAAGAGTGTACCCGAGCAACGAGGTTTCCCAAGGAGTTTTGCCAAGC 850
Db 781 CTGGGTTTCAAGAGTGTACCCGAGCAACGAGGTTTCCCAAGGAGTTTTGCCAAGC 840
Qy 911 TCCAGCGGATATCTCCCTCCAGTGTTCCTCACTGCGAGTGAAGGCGATGACCGGCGG 970
Db 901 TCCAGCGGATATCTCCCTCCAGTGTTCCTCACTGCGAGTGAAGGCGATGACCGGCGG 960
Qy 971 ATGTGGAGAGGAGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTTCATTC 1030
Db 961 ATGTGGAGAGGAGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTTCATTC 1020
Qy 1031 GGCACCAACCAATCTGTGAGGAGCGCTCATTTGACGCCCGAGTGGTCTCACTGCGGCG 1090
Db 1021 GGCACCAACCAATCTGTGAGGAGCGCTCATTTGACGCCCGAGTGGTCTCACTGCGGCG 1080
Qy 1091 CACTGCTTCTTGTGACCCCGGAGAGAGTCTTGAAGGAGTGAAGTGTACGCGGCGAC 1150
Db 1081 CACTGCTTCTTGTGACCCCGGAGAGAGTCTTGAAGGAGTGAAGTGTACGCGGCGAC 1140
Qy 1151 AGCAACTGCAACAGTTGCTGAGGAGGCTCTTCCATTTGCGAGATCATTAACAGCAAT 1210
Db 1141 AGCAACTGCAACAGTTGCTGAGGAGGCTCTTCCATTTGCGAGATCATTAACAGCAAT 1200
Qy 1211 TACACGATGAGAGAGAGCACTATGATGAGGCGCTCATGCGGCTGTCCAGGCCCTGAGC 1270
Db 1201 TACACGATGAGAGAGAGCACTATGATGAGGCGCTCATGCGGCTGTCCAGGCCCTGAGC 1260
Qy 1271 CTGTCCGCTCATCACTCCCTGCTTGTGCTCCCATGATGAGAGCACTTTAGCTCAAT 1330
Db 1261 CTGTCCGCTCATCACTCCCTGCTTGTGCTCCCATGATGAGAGCACTTTAGCTCAAT 1320
Qy 1331 GAGACTGTGATCAAGGCTTTGGCAAGACAGGAGACAGTGAACAGACATCCCCC 1390
Db 1321 GAGACTGTGATCAAGGCTTTGGCAAGACAGGAGACAGTGAACAGACATCCCCC 1380
Qy 1391 TTCTCCGGGAGGAGGATGATCTGATGACTTCAAGAAATGCAATGACTTATGCTGTC 1450
Db 1381 TTCTCCGGGAGGAGGATGATCTGATGACTTCAAGAAATGCAATGACTTATGCTGTC 1440
Qy 1451 TATGACAGTTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGTGGGGGAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGTGGGGGAGAGAC 1500
Qy 1511 TCTGCGAGGAGAGCAGCGGGGGCTTGTGTCTGTAGCAAGAAACAACCGTGTGACTG 1570
Db 1501 TCTGCGAGGAGAGCAGCGGGGGCTTGTGTCTGTAGCAAGAAACAACCGTGTGACTG 1560
Qy 1571 GAGAGTGTCAACGAGTGGGAGCAGGCTGTGGCCAGAGAAACAACCGTGTGACTGACC 1630
Db 1561 GAGAGTGTCAACGAGTGGGAGCAGGCTGTGGCCAGAGAAACAACCGTGTGACTGACC 1620
Qy 1631 AAAGTGAAGAAAGTTCTTCCCTGTGATTTAAGCAAGATGAGAGCGAGTTCAGATTGAG 1690
Db 1621 AAAGTGAAGAAAGTTCTTCCCTGTGATTTAAGCAAGATGAGAGCGAGTTCAGATTGAG 1680
Qy 1691 AAATCTTA 1699
Db 1681 AAATCTTA 1689

```
RESULT 7
US-10-353-690-99
/ Sequence 99, Application US/10353690
/ Publication No. US20030215840A1
/ GENERAL INFORMATION:
/ APPLICANT: Logan, Thomas Joseph
/ APPLICANT: Chun, Miyoung
/ APPLICANT: Galvin, Katherine M.
/ APPLICANT: Healy, Aileen
/ APPLICANT: Accorn, Susan L.
/ APPLICANT: Donoghue, Mary
/ APPLICANT: Stegaliano, Nancy
/ APPLICANT: Perodine, Jacqueline
/ APPLICANT: Rodrigue-Way, Amelie
/ TITLE OF INVENTION: Methods and compositions for treating
/ TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
/ TITLE OF INVENTION: 29002, 32126, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
/ TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
/ TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
/ TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
/ TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
/ TITLE OF INVENTION: 51130, 19469, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
/ FILE REFERENCE: MP102-018P1RONANIM
/ CURRENT APPLICATION NUMBER: US/10/353,690
/ PRIOR FILING DATE: 2003-01-29
/ PRIOR APPLICATION NUMBER: 60/353,224
/ PRIOR FILING DATE: 2002-02-01
/ PRIOR APPLICATION NUMBER: 60/364,529
/ PRIOR FILING DATE: 2002-03-15
/ PRIOR APPLICATION NUMBER: 60/373,861
/ PRIOR FILING DATE: 2002-04-19
/ PRIOR APPLICATION NUMBER: 60/376,287
/ PRIOR FILING DATE: 2002-04-29
/ PRIOR APPLICATION NUMBER: 60/388,080
/ PRIOR FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: 60/390,971
/ PRIOR FILING DATE: 2002-06-24
/ PRIOR APPLICATION NUMBER: 60/394,130
/ PRIOR FILING DATE: 2002-07-03
/ PRIOR APPLICATION NUMBER: 60/394,797
/ PRIOR FILING DATE: 2002-07-10
/ PRIOR APPLICATION NUMBER: 60/404,904
/ PRIOR FILING DATE: 2002-08-21
/ PRIOR APPLICATION NUMBER: 60/405,450
/ PRIOR FILING DATE: 2002-08-23
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 126
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 99
/ LENGTH: 2393
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-10-353-690-99
Query Match 95.6%; Score 1670.4; DB 7; Length 2393;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 258 CCAGGCAATCTCCAGCTGGTACCTCCAGGCGGGGCAATCTCCAGGCGGGGATCTCCAGC 317
Qy 241 CCAGGCAATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGC 300
Db 318 CCAGGCAATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGC 377
Qy 301 CAGGTCATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGC 360
Db 378 CAGGTCATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGC 437
Qy 361 TAGAGCAACACGAGTGGGGGCTGTACCCATCCATCTCTCCAGGTCAGACACAGC 420
Db 438 TAGAGCAACACGAGTGGGGGCTGTACCCATCCATCTCTCCAGGTCAGACACAGC 497
Qy 421 AACGAGGCGCCACGAGGAGGAGCCAGGTCAGAGCTCCCAAGTTCACTCCAGGCGGGAGG 480
Db 498 AACGAGGCGCCACGAGGAGGAGCCAGGTCAGAGCTCCCAAGTTCACTCCAGGCGGGAGG 557
Qy 481 CCAGAGCAGCTACCGCTCATCGGGTGGGCTCTCTCTCATTTGCGCTGGTGGTTCGCT 540
Db 558 CCAGAGCAGCTACCGCTCATCGGGTGGGCTCTCTCTCATTTGCGCTGGTGGTTCGCT 617
Qy 541 CATATCTCTCTCCAGTTCTGGCAGGCGCCACAGGATCAGGTACAGAGAGAGAGGA 600
Db 618 CATATCTCTCTCCAGTTCTGGCAGGCGCCACAGGATCAGGACAGAGAGAGAGGA 677
Qy 601 GAGGTGCCAGAGACCGCTGTGGCTGTGACGGGGTGGTGGACAGAGCTGAAAGAGGA 660
Db 678 GAGGTGCCAGAGACCGCTGTGGCTGTGACGGGGTGGTGGACAGAGCTGAAAGAGGA 737
Qy 661 CGAGCTGGGCTGCTGAGGTTTGAAGTGGAGCAAGTCTGCTTAAATCTACTCTGGGTC 720
Db 738 CGAGCTGGGCTGCTGAGGTTTGAAGTGGAGCAAGTCTGCTTAAATCTACTCTGGGTC 797
Qy 721 CTCCATCAGTGGCTTCCATCTGTAGCAGCACTGAAATGACTCTTCACTCAAGAAAGAC 780
Db 798 CTCCATCAGTGGCTTCCATCTGTAGCAGCACTGAAATGACTCTTCACTCAAGAAAGAC 857
Qy 781 CTGCGACAGCTGGGTTTCCAGAGTGTCAACCGGACAAACGAGGTTGCCACAGGATTT 840
Db 858 CTGCGACAGCTGGGTTTCCAGAGTGTCAACCGGACAAACGAGGTTGCCACAGGATTT 917
Qy 841 TGCCAGAGCTTCTCATCTTTGAGATACAACTCCACATCCAGAGAAAGCTCTCCAGGTC 900
Db 918 TGCCAGAGCTTCTCATCTTTGAGATACAACTCCACATCCAGAGAAAGCTCTCCAGGTC 977
Qy 901 TGAATGCCCTTCCAGCGGTATATCTCCCTCAAGTCTTCCATCTGCGAGCTGAGGGCCAT 960
Db 978 TGAATGCCCTTCCAGCGGTATATCTCCCTCAAGTCTTCCATCTGCGAGCTGAGGGCCAT 1037
Qy 961 GACCGGGCGGATCGTGGAGAGGGGCGCTGGCTGGATAGCAAGTGGCTTGGCAAGTAG 1020
Db 1038 GACCGGGCGGATCGTGGAGAGGGGCGCTGGCTGGATAGCAAGTGGCTTGGCAAGTAG 1097
Qy 1021 TCTGCACTTGGGACCAACCAATCTGTGAGGAGCGCTCATTTGACCGCCAGTGGTGTCT 1080
Db 1098 TCTGCACTTGGGACCAACCAATCTGTGAGGAGCGCTCATTTGACCGCCAGTGGTGTCT 1157
Qy 1081 CACTGCGGCCACTGCTTTCTTGTGACCCGGAGAGAGTCTTGAAGGGCTGGAAGGTGA 1140
Db 1158 CACTGCGGCCACTGCTTTCTTGTGACCCGGAGAGAGTCTTGAAGGGCTGGAAGGTGA 1217
Qy 1141 CGGGGAGACAGCAACCTGACCAAGTGGCTGAGAGGCGCTCATTTGCCAGATCATCAT 1200
Db 1218 CGGGGAGACAGCAACCTGACCAAGTGGCTGAGAGGCGCTCATTTGCCAGATCATCAT 1277
Qy 1201 CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1278 CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1337
Qy 1261 GCGCTTGAACCTGTGGCTCAATCAACCTGTGGCTTCCATTCATGAGAGAGAGCTTT 1320
```

| | | | |
|----|------|---|------|
| Db | 1338 | GCCCCGACCCCTGTCCGCTCAGATCCACCTCGCTTGGCTCCCATGTAGTACAGACACTT | 1397 |
| Qy | 1321 | TAGCCTCAATGAGACTGTCTGGATCAGAGCTTTGGCAAGCAAGGAGACAGATGACAA | 1380 |
| Db | 1398 | TAGCCTCAATGAGACTGTCTGGATCAGAGCTTTGGCAAGCAAGGAGACAGATGACAA | 1457 |
| Qy | 1381 | GACATCCCCCTTCTCCGGGAGGTGCAGGTCAATCTCATGCATTTCAAGAAATGCATGA | 1440 |
| Db | 1458 | GACATCCCCCTTCTCCGGGAGGTGCAGGTCAATCTCATGCATTTCAAGAAATGCATGA | 1517 |
| Qy | 1441 | CTACTTGGTCAATGACAGTTACCTTACCCCAAGATGATGTCTGGGGACCTTCGNG | 1500 |
| Db | 1518 | CTACTTGGTCAATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGGACCTTCGNG | 1577 |
| Qy | 1501 | GGCAGAGACTCTCCGACAGGAGACAGCGGGGGCTTTGTCTGTGACGAAACAACG | 1560 |
| Db | 1578 | GGCAGAGACTCTCCGACAGGAGACAGCGGGGGCTTTGTCTGTGACGAAACAACG | 1637 |
| Qy | 1561 | CTGCTACTGGCAGGTGTCAACAGCTGGGGCAACAGCTGTGGCCAGAGAAACAACCTGG | 1620 |
| Db | 1638 | CTGCTACTGGCAGGTGTCAACAGCTGGGGCAACAGCTGTGGCCAGAGAAACAACCTGG | 1697 |
| Qy | 1621 | TGCTTACCAACAATGACAGAAGTTCTTCCCTGATTTACAGCAAGAATGGAG | 1672 |
| Db | 1698 | TGCTTACCAACAATGACAGAAGTTCTTCCCTGATTTACAGCAAGAATGGAG | 1749 |

```

RESULT 8
US-10-156-214A-25
/ Sequence 25, Application US/10156214A
/ Publication No. US20040001801A1
/ GENERAL INFORMATION:
/ APPLICANT: Edwin L. Madison
/ APPLICANT: Joseph Edward Semple
/ APPLICANT: George P. Vlasuk
/ APPLICANT: Scott Jeffrey Kemp
/ APPLICANT: Mallareddy Komandla
/ APPLICANT: Daniel Vanna Siev
/ TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
/ TITLE OF INVENTION: Thereof
/ FILE REFERENCE: 24745-1611
/ CURRENT APPLICATION NUMBER: US/10/156,214A
/ CURRENT FILING DATE: 2002-05-23
/ NUMBER OF SEQ ID NOS: 611
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 25
/ LENGTH: 2067
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(2067)
/ OTHER INFORMATION: Nucleic acid encoding (endotheliasase 2-L) protein
US-10-156-214A-25

```

| | | | | |
|-----------------------|--------------|--------------|------------|-------------|
| Query Match | 95.2% | Score 1663.4 | DB 7 | Length 2067 |
| Best Local Similarity | 99.9% | Pred. No. 0 | | |
| Matches 1664 | Conservative | 0 | Mismatches | 1 |
| | | | Indels | 0 |
| | | | Gaps | 0 |

| | | | |
|----|-----|---|-----|
| OY | 11 | ATGAGAGGGACAGCCACCGGAATGCATCTCCAGCAMAACAACCTTCAGTGTGAGCATCT | 70 |
| Dd | 1 | ATGAGAGGGACAGCCACCGGAATGCATCTCCAGCAMAACAACCTTCAGTGTGAGCATCT | 60 |
| OY | 71 | CCAGGCCAGGCATCTCCAGCTGGGACACTCCAGGCCGGGCATCTCCAGCCCAAGCATCT | 130 |
| Dd | 61 | CCAGGCCAGGCATCTCCAGCTGGGACACTCCAGGCCGGGCATCTCCAGCCCAAGCATCT | 120 |
| OY | 131 | CCAGCCCAGGCATCTCCAGCTGGGACACTCCGGSCCGGSCATCTCCAGCCCAAGCATCT | 190 |
| Dd | 121 | CCAGCCCAGGCATCTCCAGCTGGGACACTCCGGSCCGGSCATCTCCAGCCCAAGCATCT | 180 |
| OY | 191 | CCAGCTGTACACTCCAGGCCGGGCATCTCCAGGCCGGGCATCTCCAGCCCAAGCATCT | 250 |

| | | | |
|----|------|--|------|
| Db | 181 | CCACCTGGTACACCTCCAGAGCCGGGGCATCTCCAGGCGGGGGATCTCCAGCCAGGACCT | 240 |
| Qy | 251 | CCAGCCCGGGGCATCTCCGGCTCTGGGCATCACTTTCCAGGTCCTCATCCGGAGGTATCA | 310 |
| Db | 241 | CCAGCCCGGGGCATCTCCGGCTCTGGGCATCACTTTCCAGGTCCTCATCCGGAGGTATCA | 300 |
| Qy | 311 | TTCCGCCAGGTCAAGCTCTGGGTGACAACTCCCCAAACGAGTGTACTTTTGAAGCAACA | 370 |
| Db | 301 | TTCCGCCAGGTCAAGCTCTGGGTGACAACTCCCCAAACGAGTGTACTTTTGAAGCAACA | 360 |
| Qy | 371 | CCAGTGGAGGGCTGTACACCATCTCCGATCATCTCTCCAGCTCAGCACAGCAACGAGGCC | 430 |
| Db | 361 | CCAGTGGAGGGCTGTACACCATCTCCGATCATCTCTCCAGCTCAGCACAGCAACGAGGCC | 420 |
| Qy | 431 | ACCAGGAGAGCCCGAGGTACGAGCTTGCCCAAGTTCACTTGGCGGAGGGCCAGAAACAG | 490 |
| Db | 421 | ACCAGGAGAGCCCGAGGTACGAGCTTGCCCAAGTTCACTTGGCGGAGGGCCAGAAACAG | 480 |
| Qy | 491 | CTACCCGCTCATCCGGGTGCGTGTCTCTCTCTCATTTGCTTGCTGGTGTTCGCTCATATCTCTC | 550 |
| Db | 481 | CTACCCGCTCATCCGGGTGCGTGTCTCTCTCTCATTTGCTTGCTGGTGTTCGCTCATATCTCTC | 540 |
| Qy | 551 | TTTCCAGTTCTTGGCAGGGGCACACAGGGATCAGGTACAGAGAGCAGAGGAGCTGTCCC | 610 |
| Db | 541 | TTTCCAGTTCTTGGCAGGGGCACACAGGGATCAGGTACAGAGAGCAGAGGAGCTGTCCC | 600 |
| Qy | 611 | AAGCAGCGCTGTCCGCTGTGACGGGGGTGTGTGTGACCTGCAAGCTGAAGATGACGACTGGGC | 670 |
| Db | 601 | AAGCAGCGCTGTTCGCTGTGACGGGGGTGTGTGTGACCTGCAAGCTGAAGATGACGACTGGGC | 660 |
| Qy | 671 | TGCGTGAGGTTTGAAGCTGGGACAGTCTCTGTCTTAAATCTACTTGGGTCCTCCCATCAG | 730 |
| Db | 661 | TGCGTGAGGTTTGAAGCTGGGACAGTCTCTGTCTTAAATCTACTTGGGTCCTCCCATCAG | 720 |
| Qy | 731 | TGGGCTTCCCATCTGTATGAGAGCAATGGAAATGACTCTTACTCTGAGAAAGACTGCCACAG | 790 |
| Db | 721 | TGGGCTTCCCATCTGTATGAGAGCAATGGAAATGACTCTTACTCTGAGAAAGACTGCCACAG | 780 |
| Qy | 791 | CTGGGTTTTCGAGATGTGCTCACCGGACAAACGAGGTTGCCACAGGGAATTTTGCCAAAGC | 850 |
| Db | 781 | CTGGGTTTTCGAGATGTGCTCACCGGACAAACGAGGTTGCCACAGGGAATTTTGCCAAAGC | 840 |
| Qy | 851 | TTTCTCAATCTTGAATATCAACTCCACATCCAGGAAAGCTTCCACAGGTCGTGAATGCCCT | 910 |
| Db | 841 | TTTCTCAATCTTGAATATCAACTCCACATCCAGGAAAGCTTCCACAGGTCGTGAATGCCCT | 900 |
| Qy | 911 | TCCCAAGGAGTATATCTCCCTCCCAATGTTCCCACTGCGGACCTGAGAGGGCATGACCGGGCGG | 970 |
| Db | 901 | TCCCAAGGAGTATATCTCCCTCCCAATGTTCCCACTGCGGACCTGAGAGGGCATGACCGGGCGG | 960 |
| Qy | 971 | ATTCGGGAGGAGGGGCGCTGAGCTTCGGATATGCAAGTGAGCTTGTGGCAAGTGTCTCACTTC | 1030 |
| Db | 961 | ATTCGGGAGGAGGGGCGCTGAGCTTCGGATATGCAAGTGAGCTTGTGGCAAGTGTCTCACTTC | 1020 |
| Qy | 1031 | GGCACCACCCACATCTGTGAGAGCAGCTCATTTGACGCCACGATGGGGTCTCATCTGCCGCC | 1090 |
| Db | 1021 | GGCACCACCCACATCTGTGAGAGCAGCTCATTTGACGCCACGATGGGGTCTCATCTGCCGCC | 1080 |
| Qy | 1091 | CACGTGCTTCTTGTGACCCCGGAGAAAGTCTCTGAGGGCTTGGAAAGGTGTACCGGGGCACC | 1150 |
| Db | 1081 | CACGTGCTTCTTGTGACCCCGGAGAAAGTCTCTGAGGGCTTGGAAAGGTGTACCGGGGCACC | 1140 |
| Qy | 1151 | AGCAACCTGCACAGTTTGCCTGTGAGGACGCTTCCATTGCGCGAGATCATCAACAGCAAT | 1210 |
| Db | 1141 | AGCAACCTGCACAGTTTGCCTGTGAGGACGCTTCCATTGCGCGAGATCATCAACAGCAAT | 1200 |
| Qy | 1211 | TACACCGATGAGAGAGAGACTATGACATGCGCTCTCATGCGGCTGTCCAGCCCTGACC | 1270 |
| Db | 1201 | TACACCGATGAGAGAGAGACTATGACATGCGCTCTCTCATGCGGCTGTCCAGCCCTGACC | 1260 |
| Qy | 1271 | CTGTTCGGTCAATTCACCCCTGTGCTTCCCATGAGATGAGACAGACTTATAGCTCAAT | 1330 |
| Db | 1261 | CTGTTCGGTCAATTCACCCCTGTGCTTCCCATGAGATGAGACAGACTTATAGCTCAAT | 1320 |

QY 1331 GAGACCTGCTGATCACAGGCTTTGGCAAGACAGAGGAGACAGATGACAAAGACATCCCCC 1330
DB 1321 GAGACCTGCTGATCACAGGCTTTGGCAAGACAGAGGAGACAGATGACAAAGACATCCCCC 1380
QY 1391 TTCTCTCCGGAGGAGTGCAGAGTCAATCTTCATGACCTTCAAGAAATGCAATGACTTCTGCTC 1450
DB 1381 TTCTCTCCGGAGGAGTGCAGAGTCAATCTTCATGACCTTCAAGAAATGCAATGACTTCTGCTC 1440
QY 1451 TATGACAGTTACTTACCTCCCAAGATGATGTGTGTGGGACCTTGTGGGGGGAGAGAC 1510
DB 1441 TATGACAGTTACTTACCTCCCAAGATGATGTGTGTGGGACCTTGTGGGGGGAGAGAC 1500
QY 1511 TCCTGCAAGGAGACAGCGGGGGGCTCTTGTGTGTGACAAACCAACCGCTGTGACCTG 1570
DB 1501 TCCTGCAAGGAGACAGCGGGGGGCTCTTGTGTGTGACAAACCAACCGCTGTGACCTG 1560
QY 1571 GCAGGTGTCACCACTGGGGGACAGAGCTGTGTGCGGCAAGAAACCAACCTGTGTGACCC 1630
DB 1561 GCAGGTGTCACCACTGGGGGACAGAGCTGTGTGCGGCAAGAAACCAACCTGTGTGACCC 1620
QY 1631 AAAGTGACAGAAATTCTTCCCTGATTTTACAGCAAGATGAGAGC 1675
DB 1621 AAAGTGACAGAAATTCTTCCCTGATTTTACAGCAAGATGAGAGC 1665

RESULT 9

US-10-916-758-1
Sequence 1, Application US/10916758
Publication No. US20050180977A1
GENERAL INFORMATION:
APPLICANT: Nixon, Andrew
APPLICANT: Madison, Edwin L.
TITLE OF INVENTION: ENDOTHELIASE-2 LIGANDS
FILE REFERENCE: 10280-065001
CURRENT APPLICATION NUMBER: US/10/916,758
CURRENT FILING DATE: 2004-08-12
PRIOR APPLICATION NUMBER: US 60/520,164
PRIOR FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: US 60/495,005
PRIOR FILING DATE: 2003-08-14
NUMBER OF SEQ ID NOS: 113
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2067
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2064)
US-10-916-758-1

Query Match 95.2%; Score 1663.4; DB 10; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 ATGAGAGAGGACACACACGAGATGATCTTCACAGCAAGAACCTTGAAGCTGAGCATCT 70
DB 1 ATGAGAGAGGACACACACGAGATGATCTTCACAGCAAGAACCTTGAAGCTGAGCATCT 60
QY 71 CCAGCCCAAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTTCCAGCCCAAGCATCT 130
DB 61 CCAGCCCAAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTTCCAGCCCAAGCATCT 120
QY 131 CCAGCCCAAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTTCCAGCCCAAGCATCT 190
DB 121 CCAGCCCAAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTTCCAGCCCAAGCATCT 180
QY 191 CCAGCTGATACCTTCAGAGCCGGGATCTTCCAGAGCCGGGATCTTCCAGCCCAAGCATCT 250
DB 181 CCAGCTGATACCTTCAGAGCCGGGATCTTCCAGAGCCGGGATCTTCCAGCCCAAGCATCT 240
QY 251 CCAGCCCGGGATCTCCGGGCTTGGCATCACTTTCCAGGTCTTCACTCCGGGAGGTCACTCA 310

DB 241 CCAGCCCGGGATCTCCGGGCTTGGCATCACTTTCCAGGTCTTCACTCCGGGAGGTCACTCA 300
QY 311 TCAGCCAGGTACGCTGGTGAACAACCTCCCAACCAAGAGTACCTTTGTTAGAGCAAC 370
DB 301 TCAGCCAGGTACGCTGGTGAACAACCTCCCAACCAAGAGTACCTTTGTTAGAGCAAC 360
QY 371 CCAATGGAGGGGTGTAACCATCCGATCAATCTCTGCGCAGGTGACACAGCAACAGAGGCC 430
DB 361 CCAATGGAGGGGTGTAACCATCCGATCAATCTCTGCGCAGGTGACACAGCAACAGAGGCC 420
QY 431 ACCAGGAGAGCCCAAGTACAGAGCTTCCCAAGTTACCTTGGCGGAGAGGTCAGAGAG 490
DB 421 ACCAGGAGAGCCCAAGTACAGAGCTTCCCAAGTTACCTTGGCGGAGAGGTCAGAGAG 480
QY 491 CTACCGCATCGGGTGGGCTGCTCTCTCTCAATGCGCGGGGTGTTGGCTCATGCTC 550
DB 481 CTACCGCATCGGGTGGGCTGCTCTCTCTCAATGCGCGGGGTGTTGGCTCATGCTC 540
QY 551 TTCCAGTTCTGGCAGGGCCACACAGGATCAGGTACAGAGACAGAGAGAGCTGTCCC 610
DB 541 TTCCAGTTCTGGCAGGGCCACACAGGATCAGGTACAGAGACAGAGAGAGCTGTCCC 600
QY 611 AAGCAGCTGTTGCTGTGACCGGGGTGTGACCTGCAAGCTGAAGAGTACAGAGCTGGC 670
DB 601 AAGCAGCTGTTGCTGTGACCGGGGTGTGACCTGCAAGCTGAAGAGTACAGAGCTGGC 660
QY 671 TGCGTGAAGTTTGAATGAGGACAACTCTCTGTTAAATCTACTCTGGGCTCTCCCATAG 730
DB 661 TGCGTGAAGTTTGAATGAGGACAACTCTCTGTTAAATCTACTCTGGGCTCTCCCATAG 720
QY 731 TGCGTTCATCTGTAGACAGCACTGGAATGACTCTACTCAGAGAAACCTGCGACAG 790
DB 721 TGCGTTCATCTGTAGACAGCACTGGAATGACTCTACTCAGAGAAACCTGCGACAG 780
QY 791 CTGGGTTTCCAGAGTGTCTACCGGACAAACGAGTTGCCACAGAGATTTTGGCAACAGC 850
DB 781 CTGGGTTTCCAGAGTGTCTACCGGACAAACGAGTTGCCACAGAGATTTTGGCAACAGC 840
QY 851 TTCTCAATCTTGAATGATCAACTCCACCATCAGGAAAGCTCCACAGGTCTGAATGCGCT 910
DB 841 TTCTCAATCTTGAATGATCAACTCCACCATCAGGAAAGCTCCACAGGTCTGAATGCGCT 900
QY 911 TCCAGCGGATATATCTCCCTCAGTGTCCACTGCGGACTGAGAGGCACTGACCGGCGG 970
DB 901 TCCAGCGGATATATCTCCCTCAGTGTCCACTGCGGACTGAGAGGCACTGACCGGCGG 960
QY 971 ATCGTGGAGAGGGGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGTGCACTTC 1030
DB 961 ATCGTGGAGAGGGGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGTGCACTTC 1020
QY 1031 GGACACACCCACATCTGTGAGGAGCGCTCATTTGACGCCAGTGGGTGCTCACTGCGGCC 1090
DB 1021 GGACACACCCACATCTGTGAGGAGCGCTCATTTGACGCCAGTGGGTGCTCACTGCGGCC 1080
QY 1091 CACTGCTTCTTGTGACCCGGAGAAAGTCTTGGAGGGCTGGAAGGTGTACGCGGACCC 1150
DB 1081 CACTGCTTCTTGTGACCCGGAGAAAGTCTTGGAGGGCTGGAAGGTGTACGCGGACCC 1140
QY 1151 AGCAACTGACACAGTTGCTGAGGACGCTTCCATTGCGAGATCATCATACAGCAAT 1210
DB 1141 AGCAACTGACACAGTTGCTGAGGACGCTTCCATTGCGAGATCATCATACAGCAAT 1200
QY 1211 TACACGATGAGAGAGAGCACTATGACATGAGCGCTTCAATGCGGCTGTCCAAAGCCCTGAC 1270
DB 1201 TACACGATGAGAGAGAGCACTATGACATGAGCGCTTCAATGCGGCTGTCCAAAGCCCTGAC 1260
QY 1271 CTGTCCGCTCATCATCAACCTGCTTGCCTCCCATGAGATGAGACAGACCTTTAGCTCAAT 1330
DB 1261 CTGTCCGCTCATCATCAACCTGCTTGCCTCCCATGAGATGAGACAGACCTTTAGCTCAAT 1320
QY 1331 GAGACCTGCTGATCAAGGCTTTGGCAAGACCAAGGAGACAGATGACAAAGACATCCCCC 1390

Db 1321 GAGACGCTGATGATACAGAGCTTTGGCAAGACAGGAGACAGATGACAAAGATCCCCC 1380
Qy 1391 TTCTCCGGAGGTGAGGTGAATCTCATGACTTAAAGAAATGCAATGATCTAGTCTC 1450
Db 1381 TTCTCCGGAGGTGAGGTGAATCTCATGACTTAAAGAAATGCAATGATCTAGTCTC 1440
Qy 1451 TATGACATTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGGGGGAGAGAC 1510
Db 1441 TATGACATTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGGGGGAGAGAC 1500
Qy 1511 TCCTCCAGGAGAGACAGCGGGGGCTTCTGTGTGTGAGAGAAACACCGCTGTACTG 1570
Db 1501 TCCTCCAGGAGAGACAGCGGGGGCTTCTGTGTGTGAGAGAAACACCGCTGTACTG 1560
Qy 1571 GCAGGTGTACACCTGAGGAGACAGGCTGTGTGAGAGAAACAAACCTGTGTGTACACC 1630
Db 1561 GCAGGTGTACACCTGAGGAGACAGGCTGTGTGAGAGAAACAAACCTGTGTGTACACC 1620
Qy 1631 AAAGTGACAGAGTTCTTCTCGATTTTACAGCAAGATGAGAGC 1675
Db 1621 AAAGTGACAGAGTTCTTCTCGATTTTACAGCAAGATGAGAGC 1665

RESULT 10
US-10-428-275-103
; Sequence 103, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ. ID NOS: 450
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 103
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112) .. (1869)
US-10-428-275-103

Query March 94.1%; Score 1645.4; DB 8; Length 2432;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 1; Indels 15; Gaps 1;

Qy 1 CTCAGAGACCATGAGAGGAGACAGCCAGGAATGATCTTCACAGCAAGAACACCTTACG 60
Db 102 CTCAGAGACCATGAGAGGAGACAGCCAGGAATGATCTTCACAGCAAGAACACCTTACG 161
Qy 61 TGGAGCATCTTCCAGCCAGGAGATCTCCAGCTGGAGACCTCCAGCCGGGAGATCTCCAGC 120
Db 162 TGGAGCATCTTCCAGCCAGGAGATCTCCAGCTGGAGACCTCCAGCCGGGAGATCTCCAGC 221
Qy 121 CCAGGCATCTCCAGCCAGGAGATCTCCAGCTGGAGACCTCCGGGCGGGGATCTCCAGC 180

Db 222 CCAGGCATCTCCAGCCAGGAGATCTCCAGCTGGAGACACCTCCGGGCGGGGATCTCCAGC 281
Qy 181 CCAGGCATCTCCAGCTGGATACCTCCAGCCGGGAGATCTCCAGCCG----- 228
Db 282 CCAGGCATCTCCAGCTGGATACCTCCAGCCGGGAGATCTCCAGCCGGGAGATCTCCAGC 341
Qy 229 ---GGCATCTCCAGCCAGGAGATCTCCAGCCGGGAGATCTCCAGCTGGAGACCTTTC 285
Db 342 CCAGGCATCTCCAGCCAGGAGATCTCCAGCCGGGAGATCTCCAGCTGGAGACCTTTC 401
Qy 286 CAGGTCTCATCCGGAGGATCATCATCCGCAAGTCAAGCTTGGTGAACCTTCCCAAC 345
Db 402 CAGGTCTCATCCGGAGGATCATCATCCGCAAGTCAAGCTTGGTGAACCTTCCCAAC 461
Qy 346 CAGAGTACTCTTGTAGAGACACCAAGTGGGGGCTGTACCTCCGATCATCTCTGCG 405
Db 462 CAGAGTACTCTTGTAGAGACACCAAGTGGGGGCTGTACCTCCGATCATCTCTGCG 521
Qy 406 CAGGTGACACACAGCAACAGGGGACACAGGAGAGCCAGGTACAGAGCTTGCACAGTT 465
Db 522 CAGGTGACACACAGCAACAGGGGACACAGGAGAGCCAGGTACAGAGCTTGCACAGTT 581
Qy 466 CACCTGGCGGAGAGGCGGAGAGACAGTACCGCTCATCGGTGCGTCTCTCATTTGC 525
Db 582 CACCTGGCGGAGAGGCGGAGAGAGTACCGCTCATCGGTGCGTCTCTCATTTGC 641
Qy 526 CTGTGTGTTTCTGTCATCATCTCTTTCAGTTTGGCAGAGGCGCACAGGAGATCAGTA 585
Db 642 CTGTGTGTTTCTGTCATCATCTCTTTCAGTTTGGCAGAGGCGCACAGGAGATCAGTA 701
Qy 586 CAAGGAGCAGAGGAGAGGCTGTCCAGAGCAGCTTCCGCTGTGAGCGGGGCTGTGAGCTG 645
Db 702 CAAGGAGCAGAGGAGAGGCTGTCCAGAGCAGCTTCCGCTGTGAGCGGGGCTGTGAGCTG 761
Qy 646 CAAGCTGAAGAGTACAGAGCTGGGCTGAGAGTTTGAAGTGGAGCAAGTCTTGTAA 705
Db 762 CAAGCTGAAGAGTACAGAGCTGGGCTGAGAGTTTGAAGTGGAGCAAGTCTTGTAA 821
Qy 706 AATTTACTCTGGGCTCTCCATCATGTGGCTTCCATCTGTAGAGCAACTGGAATGACTC 765
Db 822 AATTTACTCTGGGCTCTCCATCATGTGGCTTCCATCTGTAGAGCAACTGGAATGACTC 881
Qy 766 CTACTGAGAGAGCTGGCAGAGCTGGGTTTGAAGTGGAGTCAACCGGAGCAACCGAGGT 825
Db 882 CTACTGAGAGAGCTGGCAGAGCTGGGTTTGAAGTGGAGTCAACCGGAGCAACCGAGGT 941
Qy 826 TGCCCAAGGAGATTTTCCAAAGCTTCTCAATCTTGAATACAACTCCATCCAGGA 885
Db 942 TGCCCAAGGAGATTTTCCAAAGCTTCTCAATCTTGAATACAACTCCATCCAGGA 1001
Qy 886 AAGCTTCCACAGGTCTGAATGCTTCTCCAGCGGTATATCTCCTCCAGTGTCCACTG 945
Db 1002 AAGCTTCCACAGGTCTGAATGCTTCTCCAGCGGTATATCTCCTCCAGTGTCCACTG 1061
Qy 946 CGAGCTGAGGAGGAGTACCGGGGAGATGTGGAGAGGGGCGTGGCTCGGATACCAAGTG 1005
Db 1062 CGAGCTGAGGAGGAGTACCGGGGAGATGTGGAGAGGGGCGTGGCTCGGATACCAAGTG 1121
Qy 1006 GCTTTGGCAAGTGAATGTGACTTCCGACACCCACATCTGTGAGAGGAGCAGCTCATTTGA 1065
Db 1122 GCTTTGGCAAGTGAATGTGACTTCCGACACCCACATCTGTGAGAGGAGCAGCTCATTTGA 1181
Qy 1066 CGCCCAAGTGGGTCTCACTGCGGCCACTGTCTTGTGTGACCCGGAGAAAGTCTCTGGA 1125
Db 1182 CGCCCAAGTGGGTCTCACTGCGGCCACTGTCTTGTGTGACCCGGAGAAAGTCTCTGGA 1241
Qy 1126 GGGCTGGAAGGTGAGCGGGGACACAGCAACTGTCACCAAGTGTGAGGAGAGGCTTCAT 1185
Db 1242 GGGCTGGAAGGTGAGCGGGGACACAGCAACTGTCACCAAGTGTGAGGAGAGGCTTCAT 1301
Qy 1186 TGCCGAGATCATCATCAAGCAATTACCGATGAGAGAGACGATATGACATCGCCT 1245

Db 1302 TGCCGAGATCATCAACAGCAATTAACCGATGAGGAGACGATATGACATCGCCCT 1361
Qy 1246 CATCGGCTGTCAAGCCCTGACCCCTGCTCGCTCAATCCACCCCTGCTGCCCAT 1305
Db 1362 CATCGGCTGTCAAGCCCTGACCCCTGCTCGCTCAATCCACCCCTGCTGCCCAT 1421
Qy 1306 GCATGACAGACCTTTAGCTTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACAG 1365
Db 1422 GCATGACAGACCTTTAGCTTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACAG 1481
Qy 1366 GGAGACAGATGACAAAGACATCCCTTCTCCGAGAGGTTGAGGTCATCTCAATGACTT 1425
Db 1482 GGAGACAGATGACAAAGACATCCCTTCTCCGAGAGGTTGAGGTCATCTCAATGACTT 1541
Qy 1426 CAAGAATGCAATGACTTCTGATGACATGACTTACCTTCCCAAGATGATGTC 1485
Db 1542 CAAGAATGCAATGACTTCTGATGACATGACTTACCTTCCCAAGATGATGTC 1601
Qy 1486 TGGGACCTTGTGGGGGAGAGACTCTGCGAGGAGACAGCGGGGGCTTTGTCTG 1545
Db 1602 TGGGACCTTGTGGGGGAGAGACTCTGCGAGGAGACAGCGGGGGCTTTGTCTG 1661
Qy 1546 TGACGAGAAACAACCGCTGTGACCTGCGAGGTCACCCAGCTGGGGACAGGCTGTGGCA 1605
Db 1662 TGACGAGAAACAACCGCTGTGACCTGCGAGGTCACCCAGCTGGGGACAGGCTGTGGCA 1721
Qy 1606 GAGAAACAACCTGTGTGTACACCAAGTGAAGAGTTCTTCCCTGATTTACAGCA 1665
Db 1722 GAGAAACAACCTGTGTGTGTACACCAAGTGAAGAGTTCTTCCCTGATTTACAGCA 1781
Qy 1666 GATGAG 1672
Db 1782 GATGAG 1788

RESULT 11

US-10-428-275-139

Sequence 139, Application US/10428275

Publication No. US20040067505A1

GENERAL INFORMATION:

APPLICANT: Alvarez et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-585

CURRENT APPLICATION NUMBER: US/10/428, 275

CURRENT FILING DATE: 2003-05-01

PRIOR APPLICATION NUMBER: 09/966545

PRIOR FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: 09/544511

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: 60/128514

PRIOR FILING DATE: 1999-04-09

PRIOR APPLICATION NUMBER: 09/569269

PRIOR FILING DATE: 2000-05-11

PRIOR APPLICATION NUMBER: 60/134315

PRIOR FILING DATE: 1999-05-14

PRIOR APPLICATION NUMBER: 09/619252

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/789390

PRIOR FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/185548

PRIOR FILING DATE: 2000-02-25

NUMBER OF SEQ ID NOS: 450

SOFTWARE: Cureseq1st version 0.1

SEQ ID NO 139

LENGTH: 2432

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (112) .. (1869)

US-10-428-275-139

Query Match 94.0%; Score 1643.8; DB 8; Length 2432;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 CTCAGAGACCATGAGAGGGGACAGCCACGGGAATGCAATCTCCACCAAGAAACCTTCAAC 60
Db 102 CTCAGAGACCATGAGAGGGGACAGCCACGGGAATGCAATCTCCACCAAGAAACCTTCAAC 161
Qy 61 TGAGACATCTCCAGACCCAGGATCTCCAGCTGGGACACCTCCAGGCGGATCTCCAGC 120
Db 162 TGAGACATCTCCAGACCCAGGATCTCCAGCTGGGACACCTCCAGGCGGATCTCCAGC 221
Qy 121 CCAGGATCTCCAGACCCAGGATCTCCAGCTGGGACACCTCCAGGCGGATCTCCAGC 180
Db 222 CCAGGATCTCCAGACCCAGGATCTCCAGCTGGGACACCTCCAGGCGGATCTCCAGC 281
Qy 181 CCAGGATCTCCAGCTGGTACCTCCAGGCGGATCTCCAGGCGGATCTCCAGC 228
Db 282 CCAGGATCTCCAGCTGGTACCTCCAGGCGGATCTCCAGGCGGATCTCCAGC 341
Qy 229 ---GGCATCTCCAGACCCAGGATCTCCAGCTGGGACATCTCCAGCTGGATCACTTTC 285
Db 342 CCAGGATCTCCAGACCCAGGATCTCCAGCTGGGACATCTCCAGCTGGATCACTTTC 401
Qy 286 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 345
Db 402 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 461
Qy 346 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 405
Db 462 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 521
Qy 406 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 465
Db 522 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 581
Qy 466 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 525
Db 582 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 641
Qy 526 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 585
Db 642 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 701
Qy 586 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 645
Db 702 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 761
Qy 646 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 705
Db 762 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 821
Qy 706 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 765
Db 822 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 881
Qy 766 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 825
Db 882 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 941
Qy 826 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 885
Db 942 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 1001
Qy 886 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 945
Db 1002 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 1061
Qy 946 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 1005
Db 1062 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 1121
Qy 1006 CAGGTCTCATCGGACAGTTCATATCCGCAAGTCAAGCTCGGTGACAACTTCCCAAC 1065

1122 GCCTTGCAAGTAGTGTGCACTTGGCAGCACCCACCATCTGTGAGGACGCTCATTTGA 1181
1066 CGCCCAATGGGTGTCTCACTGCGGCCCACTGTTTGTGTACCCGGGAGAGAGGCTCGGA 1125
1182 CGCCCAATGGGTGTCTCACTGCGGCCCACTGTTTGTGTACCCGGGAGAGAGGCTCGGA 1241
1126 GGGCTGGAAGGTGTGACCGGGACACGACCAACCTTGCACAGTTGGCTTGAAGGAGCTTCAT 1185
1242 GGGCTGGAAGGTGTGACCGGGACACGACCAACCTTGCACAGTTGGCTTGAAGGAGCTTCAT 1301
1186 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGAGAGAGCATATGACATGCGCTT 1245
1302 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGAGAGAGCATATGACATGCGCTT 1361
1246 CATGGCTGTCCCAAGCCCTGAGCCCTGTGCTGCTCAATCCACCTTGGCTTGGCTTCCAT 1305
1362 CATGGCTGTCCCAAGCCCTGAGCCCTGTGCTGCTCAATCCACCTTGGCTTGGCTTCCAT 1421
1306 GCATGACAGACCTTTGACCTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACGAG 1365
1422 GCATGACAGACCTTTGACCTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACGAG 1481
1366 GGAAGCAGATGACAGACATCCCTTCTTCCGGAGGTGAGGTCAATCATGCACTT 1425
1482 GGAAGCAGATGACAGACATCCCTTCTTCCGGAGGTGAGGTCAATCATGCACTT 1541
1426 CAAGAATGCAATGACTTGTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTC 1485
1542 CAAGAATGCAATGACTTGTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTC 1601
1486 TGGGAGCCTTGTGAGGGGAGAGAGCTCTGTCAGGAGAGAGAGGGGGGCTTGTCTG 1545
1602 TGGGAGCCTTGTGAGGGGAGAGAGCTCTGTCAGGAGAGAGAGGGGGGCTTGTCTG 1661
1546 TGAGCAGAACCAACCGCTGTGTACTGTGAGAGGTGTACAGGCTGGGCA 1605
1662 TGAGCAGAACCAACCGCTGTGTACTGTGAGAGGTGTACAGGCTGGGCA 1721
1606 GAGAAACAAACCTGTGTGTATACCAAAAGTGTACAGAAAGTTTCTTCCGTGATTTACAGCA 1665
1722 GAGAAACAAACCTGTGTGTATACCAAAAGTGTACAGAAAGTTTCTTCCGTGATTTACAGCA 1781
1666 GATGGAG 1672
1782 GATGGAG 1788

RESULT 12
US-10-428-275-141
Sequence 141, Application US/10428275
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428, 275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25

NUMBER OF SEQ ID NOS: 450
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 141
LENGTH: 2432
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (112) .. (1869)
US-10-428-275-141

Query Match 94.0%; Score 1643.8; DB 8; Length 2432;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

1 CTGAGAGACCATGAGAGGAGACAGCCAGGAATGATCTTCAGAGAAACACTTTCAGC 60
102 CTCAGAGACCATGAGAGGAGACAGCCAGGAATGATCTTCAGAGAAACACTTTCAGC 161
61 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 120
162 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 221
121 CCAGGCACTTCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 180
222 CCAGGCACTTCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 281
181 CCAGGCACTTCAGCTGTACCTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGC 228
282 CCAGGCACTTCAGCTGTACCTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGC 341
229 ---GGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 385
342 CCAGGCACTTCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 401
286 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 345
402 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 461
346 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 405
462 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 521
406 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 465
522 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 581
466 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 525
582 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 641
526 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 585
642 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 701
586 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 645
702 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 761
646 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 705
762 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 821
706 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 765
822 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 881
766 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 825
882 CAGGCTCTCATCCGAGGATCTCCAGCTGGGAGACCTCCAGGCGGGCATCTCCAGC 941
826 TGGCAGAGGATTTTCCAAAGCTTCTCAATTTGATGATCAACTTCACATTCAGGA 885

Db 942 TGCCACAGGATTTTGGCAACAGCTTCTCAATCTTGAGATACCACTCCACCATCCAGGA 1001
Qy 886 AAGCTCCACAGGTCTGAATGCTCCCTCCAGAGGATATCTCCCTCCAGTGTCCACTG 945
Db 1002 AAGCTCCACAGGTCTGAATGCTCCCTCCAGAGGATATCTCCCTCCAGTGTCCACTG 1061
Qy 946 CGAAGTGAAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGCTTGATGCAAGTG 1005
Db 1062 CGAAGTGAAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGCTTGATGCAAGTG 1121
Qy 1006 GCCTTGCAAGTGAAGTCTGCACTTCGAGACCACTCTGTGAGAGGACGCTCATTTGA 1065
Db 1122 GCCTTGCAAGTGAAGTCTGCACTTCGAGACCACTCTGTGAGAGGACGCTCATTTGA 1181
Qy 1066 CGCCAGTGGTGTCTCACTGCGGCCCACTGCTTCTGTGAGACCGGAGAGAGTCTTGA 1125
Db 1182 CGCCAGTGGTGTCTCACTGCGGCCCACTGCTTCTGTGAGACCGGAGAGAGTCTTGA 1241
Qy 1126 GGGCTGGAAGGTGTACCGCGGACCAAGCACTTGACCAAGTTGCTTGAGGAGGCTTCAT 1185
Db 1242 GGGCTGGAAGGTGTACCGCGGACCAAGCACTTGACCAAGTTGCTTGAGGAGGCTTCAT 1301
Qy 1186 TGCCGAGATCATCTCAACAGCAATTAACCGATGAGAGAGAGCACTATGACATGCGCTT 1245
Db 1302 TGCCGAGATCATCTCAACAGCAATTAACCGATGAGAGAGAGCACTATGACATGCGCTT 1361
Qy 1246 CATCGGCTGTCTCAAGCGCTTGAACCTGTCCGCTCACTCACTCCCTGCTTCCCTCAT 1305
Db 1362 CATCGGCTGTCTCAAGCGCTTGAACCTGTCCGCTCACTCACTCCCTGCTTCCCTCAT 1421
Qy 1306 GCATGAGACAGCTTTAGCTCTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACGAG 1365
Db 1422 GCATGAGACAGCTTTAGCTCTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACGAG 1481
Qy 1366 GGAGACAGATGACAGACATCCCTCTCTCCGAGAGGTGAGAGTCAATCTCACTGACTT 1425
Db 1482 GGAGACAGATGACAGACATCCCTCTCTCCGAGAGGTGAGAGTCAATCTCACTGACTT 1541
Qy 1426 CAAGAAATGCAATGACTACTTGTCTATGACAGTTACTTCCCAAGAGATGATGTGTC 1485
Db 1542 CAAGAAATGCAATGACTACTTGTCTATGACAGTTACTTCCCAAGAGATGATGTGTC 1601
Qy 1486 TGGGAGCTTGTGTGGGGGAGAGACTCTCTGCGAGGAGACAGCGGGGGGCTTGTGTG 1545
Db 1602 TGGGAGCTTGTGTGGGGGAGAGACTCTCTGCGAGGAGACAGCGGGGGGCTTGTGTG 1661
Qy 1546 TGAGCAGAAACAACGGCTGTACTCTGAGAGGTGTACCAAGCTGGGGGACAGGCTGTGCGCA 1605
Db 1662 TGAGCAGAAACAACGGCTGTACTCTGAGAGGTGTACCAAGCTGGGGGACAGGCTGTGCGCA 1721
Qy 1606 GAGAAACAACCTGTGTGTGTACCAAGAGTTCTTCCCTGAGATTTACAGCAA 1665
Db 1722 GAGAAACAACCTGTGTGTGTGTACCAAGAGTTCTTCCCTGAGATTTACAGCAA 1781
Qy 1666 GATGGAG 1672
Db 1782 GATGGAG 1788

RESULT 13
US-10-428-275-143
Sequence 143, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428, 275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Cursesqlist version 0.1
SEQ ID NO 143
LENGTH: 2432
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (112)..(1347)
US-10-428-275-143

Query Match 94.0%; Score 1643.8; DB 8; Length 2432;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

1 CTCGAGACCATGAGAGGGGACAGCCAGGGAATGATCTCCAGCAAGAACCTTCAGC 60
Db 102 CTCGAGACCATGAGAGGGAGACAGCCAGGGAATGATCTCCAGCAAGAACCTTCAGC 161
Qy 61 TGGAGCATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 120
Db 162 TGGAGCATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 221
Qy 121 CCAGGATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 180
Db 222 CCAGGATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 281
Qy 181 CCAGGATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 228
Db 282 CCAGGATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 341
Qy 229 ---GGCATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 285
Db 342 CCAGGATCTTCAGACCCAGGATCTCGAGCTGGAGCACTTCAGACCGGCGATCTCAGC 401
Qy 286 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTCAAGCTTCGCTGACCACTTCCCAAC 345
Db 402 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 461
Qy 346 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 405
Db 462 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 521
Qy 406 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 465
Db 522 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 581
Qy 466 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 525
Db 582 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 641
Qy 526 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 585
Db 642 CAGGTCTCATCCGGCAGGTTCATATCCGCCAGGTTCGCTGACCACTTCCCAAC 701
Qy 586 CAGGAGCAGAGGAGAGTGTCCCAAGCAAGCTGTGTGAGCCGAGGTGTGAGCTG 645
Db 702 CAGGAGCAGAGGAGAGTGTCCCAAGCAAGCTGTGTGAGCCGAGGTGTGAGCTG 761
Qy 646 CAGGAGCAGAGGAGAGTGTCCCAAGCAAGCTGTGTGAGCCGAGGTGTGAGCTG 705

Db 762 CAAGCTGAGAGTACGAGCTGGGCTGGGTGAGTTGATCGGACCAAGTCTCTGCTTAA 821
Qy 706 AATCTACTCTGGTGTCCCTCCATCAGTGGCTTCCATCTGTAGCGACCACTGGATGACTC 765
Db 822 AATCTACTCTGGTGTCCCTCCATCAGTGGCTTCCATCTGTAGCGACCACTGGATGACTC 881
Qy 766 CTACTCAGAGAAAGCTGCGAGCTGGTGTGAGAGTCTCAACCGACCAAGAGT 825
Db 882 CTACTCAGAGAAAGCTGCGAGCTGGTGTGAGAGTCTCAACCGACCAAGAGT 941
Qy 826 TGCCCAAGAGGATTTTGGCAACAGCTTCTCAATCTTGAGATCAACTCCACATCCAGGA 885
Db 942 TGCCCAAGAGGATTTTGGCAACAGCTTCTCAATCTTGAGATCAACTCCACATCCAGGA 1001
Qy 886 AAGCTTCCACAGGTCTGAATGCTTCCACAGGATATCTCCCTCCAGTGTCCCACTG 945
Db 1002 AAGCTTCCACAGGTCTGAATGCTTCCACAGGATATCTCCCTCCAGTGTCCCACTG 1061
Qy 946 CGGACTGAGGCGCAATGACCGGGCGGATGTGGAGAGGGCGCTGGCTCGATAGCAAGTG 1005
Db 1062 CGGACTGAGGCGCAATGACCGGGCGGATGTGGAGAGGGCGCTGGCTCGATAGCAAGTG 1121
Qy 1006 GCCTTGGCAAGTGAAGTGTGCACTTCCGACCAACCACTGTGGAGCAAGCTCATTTGA 1065
Db 1122 GCCTTGGCAAGTGAAGTGTGCACTTCCGACCAACCACTGTGGAGCAAGCTCATTTGA 1181
Qy 1066 CGCCCAAGTGGGTCTCACTGCTCCGCCCACTGCTTCTTCTGTGACCGGGAGAGTCTTGA 1125
Db 1182 CGCCCAAGTGGGTCTCACTGCTCCGCCCACTGCTTCTTCTGTGACCGGGAGAGTCTTGA 1241
Qy 1126 GGGCTGAGAGGTGTACGCGGGGACACGACCACTGACAGGATGTGCTGGAGAGGCTCCAT 1185
Db 1242 GGGCTGAGAGGTGTACGCGGGGACACGACCACTGACAGGATGTGCTGGAGAGGCTCCAT 1301
Qy 1186 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGGAGGACGACTATGACATGCGCCT 1245
Db 1302 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGGAGGACGACTATGACATGCGCCT 1361
Qy 1246 CATGCGGCTGTCCAAAGCCCTGAGACCTGTCCGCTCACTCAACCTGTGCTTCCCTCCAT 1305
Db 1362 CATGCGGCTGTCCAAAGCCCTGAGACCTGTCCGCTCACTCAACCTGTGCTTCCCTCCAT 1421
Qy 1306 GCATGAGACAGCTTTAGCCTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACCG 1365
Db 1422 GCATGAGACAGCTTTAGCCTCAATGAGACCTGCTGATCAAGGCTTTGGCAAGACCG 1481
Qy 1366 GAGACAGATGACAAAGATCCCTTCTCCGGAGGAGTGCATCTCATGACTT 1425
Db 1482 GAGACAGATGACAAAGATCCCTTCTCCGGAGGAGTGCATCTCATGACTT 1541
Qy 1426 CAAGAATGCAATGACTTATGCTGTATGACAGTTACCTTACCCCAAGATGATGTGTC 1485
Db 1542 CAAGAATGCAATGACTTATGCTGTATGACAGTTACCTTACCCCAAGATGATGTGTC 1601
Qy 1486 TGGGAGCCTTGTGGGGGAGAGACTCTGCGAGGAGACAGCGGGGGGCTCTTGTGTG 1545
Db 1602 TGGGAGCCTTGTGGGGGAGAGACTCTGCGAGGAGACAGCGGGGGGCTCTTGTGTG 1661
Qy 1546 TGACGAGAAACAACGCTGTGTAAGTGTGACCAAGTGTGAGGAGCAAGGCTGTGAGCA 1605
Db 1662 TGACGAGAAACAACGCTGTGTAAGTGTGACCAAGTGTGAGGAGCAAGGCTGTGAGCA 1721
Qy 1606 GAGAAACAACCTGTGTGTACCAAAAGTGAACAGAGTTCTTCCCTGGATTTACAGCAA 1665
Db 1722 GAGAAACAACCTGTGTGTGTACCAAAAGTGAACAGAGTTCTTCCCTGGATTTACAGCAA 1781
Qy 1666 GATGAG 1672
Db 1782 GATGAG 1788

Sequence 123, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428, 275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Churaseglit version 0.1
SEQ ID NO 123
LENGTH: 1771
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (114)..(1771)
US-10-428-275-123

Query Match 93.8%; Score 1640; DB 8; Length 1771;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

Qy 8 ACATGAGAGAGGAGACCGGAGATGATCTTCCAGCAAGACACTTCAAGCTGAGACA 67
Db 11 ACATGAGAGAGGAGACCGGAGATGATCTTCCAGCAAGACACTTCAAGCTGAGACA 70
Qy 68 TCTCAGCCCAAGGATCTCCAGCTGGGACCTCCAGCCGGGATCTCCAGCCAGGCA 127
Db 71 TCTCAGCCCAAGGATCTCCAGCTGGGACCTCCAGCCGGGATCTCCAGCCAGGCA 130
Qy 128 TCTCAGCCCAAGGATCTCCAGCTGGGACCTCCAGCCGGGATCTCCAGCCAGGCA 187
Db 131 TCTCAGCCCAAGGATCTCCAGCTGGGACCTCCAGCCGGGATCTCCAGCCAGGCA 190
Qy 188 TCTCAGCTGTATCACTCCAGGCGGCGATCTCCAGGCGG-----GGCA 232
Db 191 TCTCAGCTGTATCACTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGA 250
Qy 233 TCTCAGCCCAAGGATCTCCAGCTGGGACCTCCAGCCGGGATCTCCAGCTTCCAGGTC 292
Db 251 TCTCAGCCCAAGGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGTC 310
Qy 293 TCTCAGCCCAAGGATCTCCAGCTGGGACCTCCAGCCGGGATCTCCAGCTTCCAGGTC 352
Db 311 TCTCAGCCCAAGGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGTC 370
Qy 353 TACTTGTAGAGCAACAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 412
Db 371 TACTTGTAGAGCAACAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 430
Qy 413 GCACGAGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 472
Db 431 GCACGAGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 490
Qy 473 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 532

Ds 491 CGGAGGGCCGAAAGACGACCCGCTCATCGGGTGGCTGCTCTTCATTTGCCCTGGT 550
Qy 533 GTTTCGCTCATCTCTCTTCCAGTCTGGCAGGGCCACACAGGAGATCAAGTACAAGAG 592
Ds 551 GTTTCGCTCATCTCTCTTCCAGTCTGGCAGGGCCACACAGGAGATCAAGTACAAGAG 610
Qy 593 CAGAGGAGAGCTGTCCCAAGCAGCTGTTGCTGTGA CCGGGGTGTGAGCTGCAAGCTG 652
Ds 611 CAGAGGAGAGCTGTCCCAAGCAGCTGTTGCTGTGA CCGGGGTGTGAGCTGCAAGCTG 670
Qy 653 AAGAGTACAGAGCTGGGCTGGTGAAGGTTTGA CTGGGACAAGTCTCTGCTTAAATCTAC 712
Ds 671 AAGAGTACAGAGCTGGGCTGGTGAAGGTTTGA CTGGGACAAGTCTCTGCTTAAATCTAC 730
Qy 713 TCTGGGCTCTCCCATCACTAGTGGCTTCCCATCTGTAGACGACA CTGGAATGACTCTACTCA 772
Ds 731 TCTGGGCTCTCCCATCACTAGTGGCTTCCCATCTGTAGACGACA CTGGAATGACTCTACTCA 790
Qy 773 GAGAAAGACTGCGCAGCAGCTGGGTTTGAAGTGTCTCA CCGGACAA CCGAGGTTGCCAC 832
Ds 791 GAGAAAGACTGCGCAGCAGCTGGGTTTGAAGTGTCTCA CCGGACAA CCGAGGTTGCCAC 850
Qy 833 AAGGATTTTGGCAACAGCTTCTCAATCTTGAATACAA CTCCACATCCAGAAAGCTC 892
Ds 851 AAGGATTTTGGCAACAGCTTCTCAATCTTGAATACAA CTCCACATCCAGAAAGCTC 910
Qy 893 CACAGGCTCTGAATGCTTCCACAGGGAATATCTCCCTCAGTGTCCCATCTGGCAGCTG 952
Ds 911 CACAGGCTCTGAATGCTTCCACAGGGAATATCTCCCTCAGTGTCCCATCTGGCAGCTG 970
Qy 953 AAGGCTCATGACCGGAGCGGATCTGTGAGAGGGGCGCTGGCTCTGGATAGCAATGGCTTTGG 1012
Ds 971 AAGGCTCATGACCGGAGCGGATCTGTGAGAGGGGCGCTGGCTCTGGATAGCAATGGCTTTGG 1030
Qy 1013 CAAGTGAAGTGCACCTTGAGACCA CCGACATCTGTGAGGACGAGCTCATTTGAGGCGCAG 1072
Ds 1031 CAAGTGAAGTGCACCTTGAGACCA CCGACATCTGTGAGGACGAGCTCATTTGAGGCGCAG 1090
Qy 1073 TGGGTGCTCACTGCGCGCCACCTGCTTCTTCTGTGA CCGCGGAGAAAGTCTTGAGAGGCTG 1132
Ds 1091 TGGGTGCTCACTGCGCGCCACCTGCTTCTTCTGTGA CCGCGGAGAAAGTCTTGAGAGGCTG 1150
Qy 1133 AAGGTGTAACGGGGGACCCAGCA CCGTGCACCAAGTTGCTGTAGAGCAAGCTTCATTGGCGAG 1192
Ds 1151 AAGGTGTAACGGGGGACCCAGCA CCGTGCACCAAGTTGCTGTAGAGCAAGCTTCATTGGCGAG 1210
Qy 1193 ATCATCATCAACAGCAATTTACACGATGAGAGGACGACATATGACATCGCCCTCATGCGG 1252
Ds 1211 ATCATCATCAACAGCAATTTACACGATGAGAGGACGACATATGACATCGCCCTCATGCGG 1270
Qy 1253 CTGTCCAGAGCCCTGAGCCCTGTCCGCTCA CATTCCACCTGTGCTTCCCTCCATGATGGA 1312
Ds 1271 CTGTCCAGAGCCCTGAGCCCTGTCCGCTCA CATTCCACCTGTGCTTCCCTCCATGATGGA 1330
Qy 1313 CAGACCTTTAGCCTCAATGAGACCTGTGTATCA CAGGCTTTGGACAAGACCAAGGAGACA 1372
Ds 1331 CAGACCTTTAGCCTCAATGAGACCTGTGTATCA CAGGCTTTGGACAAGACCAAGGAGACA 1390
Qy 1373 GATGACAAAGACATCCCTCTCTCCGGAAGTGAAGGTCAT TCTCATGATTTCAAGAAA 1432
Ds 1391 GATGACAAAGACATCCCTCTCTCCGGAAGTGAAGGTCAT TCTCATGATTTCAAGAAA 1450
Qy 1433 TGCATATGACTACTGTGTATGACAGTTACTTACCCCAAGATGATGTGTGCTGGGAC 1492
Ds 1451 TGCATATGACTACTGTGTATGACAGTTACTTACCCCAAGATGATGTGTGCTGGGAC 1510
Qy 1493 CTTGTGGGGGACAGACTCTGTCCAGGAGACACAGCGGGGGGCTCTTGTCTGTAGACAG 1552
Ds 1511 CTTGTGGGGGACAGACTCTGTCCAGGAGACACAGCGGGGGGCTCTTGTCTGTAGACAG 1570
Qy 1553 AACAAACGGCTGTACTCTGGCAGGTGTACCA CCGCTGGGGGACAGGCTGTGTGGCAGAGAAC 1612
Ds 1571 AACAAACGGCTGTACTCTGGCAGGTGTACCA CCGCTGTGTGGCAGAGAAC 1630

Qy 1613 AACCTGTGTGTACACCAAAAGTACAGAAATTCTTCTCGATTTTACAGCAAGATGGAG 1672
Ds 1631 AACCTGTGTGTGTACACCAAAAGTACAGAAATTCTTCTCGATTTTACAGCAAGATGGAG 1690
RESULT 15
US-10-428-275-119
/ Sequence 119, Application US/10428275
/ Publication No. US20040067505A1
/ GENERAL INFORMATION:
/ APPLICANT: Alvarez et al.
/ TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
/ FILE REFERENCE: 21402-585
/ CURRENT APPLICATION NUMBER: US/10/428, 275
/ PRIOR FILING DATE: 2003-05-01
/ PRIOR APPLICATION NUMBER: 09/966545
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 09/544511
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: 60/128514
/ PRIOR FILING DATE: 1999-04-09
/ PRIOR APPLICATION NUMBER: 09/569269
/ PRIOR FILING DATE: 2000-05-11
/ PRIOR APPLICATION NUMBER: 60/134315
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/619252
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/789390
/ PRIOR FILING DATE: 2001-02-23
/ PRIOR APPLICATION NUMBER: 60/185548
/ PRIOR FILING DATE: 2000-02-25
/ NUMBER OF SEQ ID NOS: 450
/ SOFTWARE: CuroSeqList version 0.1
/ SEQ ID NO 119
/ LENGTH: 1779
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1779)
US-10-428-275-119
Query Match 93.8%; Score 1640; DB 8; Length 1779;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 0; Indels 15; Gaps 1;
Qy 8 ACCATGAGAGGACACACACGGAATGCATCTTCCAGCAAGAACCTTTCAGCTGGAGCA 67
Ds 10 ACCATGAGAGGACACACACGGAATGCATCTTCCAGCAAGAACCTTTCAGCTGGAGCA 69
Qy 68 TCTCCAGCCCAAGGATCTCCAGCTGGAGACACTCCAGGCGGGGATCTCCAGCCAGGCA 127
Ds 70 TCTCCAGCCCAAGGATCTCCAGCTGGAGACACTCCAGGCGGGGATCTCCAGCCAGGCA 129
Qy 128 TCTCCAGCCCAAGGATCTCCAGCTGGAGACACTCCAGGCGGGGATCTCCAGCCAGGCA 187
Ds 130 TCTCCAGCCCAAGGATCTCCAGCTGGAGACACTCCAGGCGGGGATCTCCAGCCAGGCA 189
Qy 188 TCTCCAGCTGTACACTCCAGGCGGGGATCTCCAGGCGG-----GGCA 232
Ds 190 TCTCCAGCTGTACACTCCAGGCGGGGATCTCCAGGCGGCGGATCTCCAGCCAGGCA 249
Qy 233 TCTCCAGCCCAAGGATCTCCAGGCGGGGATCTCCAGGCTGTGGCATGCTTCCAGGTTCC 292
Ds 250 TCTCCAGCCCAAGGATCTCCAGGCGGGGATCTCCAGGCTGTGGCATGCTTCCAGGTTCC 309
Qy 293 TCATCCGAGAGGTATCATCCGCGAGGTGAGCTCGGTGACAACTCCCAACGAGAGTG 352
Ds 310 TCATCCGAGAGGTATCATCCGCGAGGTGAGCTCGGTGACAACTCCCAACGAGAGTG 369
Qy 353 TACCTTGTAGAGCAACACAGTGGGGGCTGTACCATCCGATATCTCTGTGCCAGGTCA 412

Db 370 TACCTTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTCCAGGTCA 429
 Qy 413 GCACAGCAACAGGGCCACAGGGAGAGCCAGGTACAGAGCTGGCCCAAGTTACCTGG 472
 Db 430 GCACAGCAACAGGGCCACAGGGAGAGCCAGGTACAGAGCTGGCCCAAGTTACCTGG 489
 Qy 473 CGGAGGGCCAGAGAGAGTACCGCTCATCGGGTGCCTCTCTCTCATTTGCCCTGGT 532
 Db 490 CGGAGGGCCAGAGAGAGTACCGCTCATCGGGTGCCTCTCTCTCATTTGCCCTGGT 549
 Qy 533 GTTTTCGTCATCATCTCTCTCAAGTTCTGGAGGGCCACAGAGGATCAGTTACAGAG 592
 Db 550 GTTTTCGTCATCATCTCTCTCAAGTTCTGGAGGGCCACAGAGGATCAGTTACAGAG 609
 Qy 593 CAGAGGAGAGAGTCTCCCAAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652
 Db 610 CAGAGGAGAGAGTCTCCCAAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
 Qy 653 AAGAGTACGAGCTGGGCTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 712
 Db 670 AAGAGTACGAGCTGGGCTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
 Qy 713 TCTGGGCTCTCCATCAGTGGCTTCCATCTGTAGCAGCAATGTAGTACTCTACTCA 772
 Db 730 TCTGGGCTCTCCATCAGTGGCTTCCATCTGTAGCAGCAATGTAGTACTCTACTCA 789
 Qy 773 GAGAGAGCTTCCAGAGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 832
 Db 790 GAGAGAGCTTCCAGAGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
 Qy 833 AAGGATTTTGGCAACAGCTTCCATCTGTAGATCAACTCCACATCCAGAGAAAGCTC 892
 Db 850 AAGGATTTTGGCAACAGCTTCCATCTGTAGATCAACTCCACATCCAGAGAAAGCTC 909
 Qy 893 CACAGTCTGATGCGCTTCCAGCGGTATATCTCCCTCAAGTGTCCCACTGGCGACTG 952
 Db 910 CACAGTCTGATGCGCTTCCAGCGGTATATCTCCCTCAAGTGTCCCACTGGCGACTG 969
 Qy 953 AAGGCTATGACCGGGCGGATCTGTGGAGGGGCGCTGGCTCGATAGCAAGTGGCTTGG 1012
 Db 970 AAGGCTATGACCGGGCGGATCTGTGGAGGGGCGCTGGCTCGATAGCAAGTGGCTTGG 1029
 Qy 1013 CAAGTGAAGTGTGCACTTGGGACCAACCAATCTGTGAGGCAAGCTATTGAGCGCCAG 1072
 Db 1030 CAAGTGAAGTGTGCACTTGGGACCAACCAATCTGTGAGGCAAGCTATTGAGCGCCAG 1089
 Qy 1073 TGGGTGCTCACTGGCGCCCACTGTCTTCTGTGTGACCCGAGAAAGTCTGTGAGGGCTGG 1132
 Db 1090 TGGGTGCTCACTGGCGCCCACTGTCTTCTGTGTGACCCGAGAAAGTCTGTGAGGGCTGG 1149
 Qy 1133 AAGGTGTAAGGGGGACCAAGCACTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 1192
 Db 1150 AAGGTGTAAGGGGGACCAAGCACTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 1209
 Qy 1193 ATCATCATCAACAGCAATTAACAGGATGAGAGGAGCACTATGACATGAGCCCTCATGCGG 1252
 Db 1210 ATCATCATCAACAGCAATTAACAGGATGAGAGGAGCACTATGACATGAGCCCTCATGCGG 1269
 Qy 1253 CTGTCCAAGCCCTGACCCCTGTGCTGCTCATTCACCCCTGTGTGCTTCCCATGTGATGA 1312
 Db 1270 CTGTCCAAGCCCTGACCCCTGTGCTGCTCATTCACCCCTGTGTGCTTCCCATGTGATGA 1329
 Qy 1313 CAGACCTTTAGGCTCATGAGACCTGTGTGATCACAGGCTTTGGCAAGACCAAGGAGACA 1372
 Db 1330 CAGACCTTTAGGCTCATGAGACCTGTGTGATCACAGGCTTTGGCAAGACCAAGGAGACA 1389
 Qy 1373 GATGCAAGAGATCCCTCTCTCGGAGGGTGCAGGTCAATCTCATGACTTCAAGAAA 1432
 Db 1390 GATGCAAGAGATCCCTCTCTCGGAGGGTGCAGGTCAATCTCATGACTTCAAGAAA 1449
 Qy 1433 TGCAATGACTACTTGTGTATGACATTAACCTTACCCCAAGAGATGTGTGTGGGAC 1492
 Db 1450 TGCAATGACTACTTGTGTATGACATTAACCTTACCCCAAGAGATGTGTGTGGGAC 1509

Search completed: September 18, 2006, 11:08:56
 Job time : 2313 secs

Qy 1493 CTTGTGGGGGACAGAGCTCTGTGCGAGAGACAGCGGGGGGCTCTTGTGTGTGAGCAG 1552
 Db 1510 CTTGTGGGGGACAGAGCTCTGTGCGAGAGACAGCGGGGGGCTCTTGTGTGTGAGCAG 1569
 Qy 1553 AACACCGCTGTACTGTGAGAGTGTACACAGCTGGGGCACAGGCTGTGGCCAGAGAAAC 1612
 Db 1570 AACACCGCTGTACTGTGAGAGTGTACACAGCTGGGGCACAGGCTGTGGCCAGAGAAAC 1629
 Qy 1613 AACCTGTGTGTACACCAAGTGTCTTCTCTGATTTACAGCAAGATGAG 1672
 Db 1630 AACCTGTGTGTACACCAAGTGTCTTCTCTGATTTACAGCAAGATGAG 1689

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 08:45:25 ; Search time 1950 Seconds
(without alignments)
1572.094 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 1748
Sequence: 1 ctcagagaccatgagagaggg.....sgctcgtcgtacgcgagaa 1748

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2330496 seqs, 876882855 residues

Total number of hits satisfying chosen parameters: 4660992

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US12_NEW_PUB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US13_NEW_PUB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US14_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-----------------------|--------------------|
| 1 | 1670.4 | 95.6 | 2393 | US-11-266-748A-22522 | Sequence 22522, A |
| 2 | 709.8 | 40.6 | 930 | US-11-266-748A-172692 | Sequence 172692, A |
| 3 | 240.2 | 13.7 | 255 | US-11-266-748A-11696 | Sequence 11696, A |
| 4 | 240.2 | 13.7 | 255 | US-11-266-748A-64397 | Sequence 64397, A |
| 5 | 240.2 | 13.7 | 255 | US-11-266-748A-67229 | Sequence 67229, A |
| 6 | 193.2 | 11.1 | 2063 | US-10-196-749-329 | Sequence 329, App |
| 7 | 193.2 | 11.1 | 2063 | US-11-101-316-111 | Sequence 111, App |
| 8 | 193.2 | 11.1 | 2063 | US-11-376-673-111 | Sequence 111, App |
| 9 | 193.2 | 11.1 | 2104 | US-11-145-307A-34 | Sequence 34, Appl |
| 10 | 182.8 | 10.5 | 2412 | US-10-196-749-63 | Sequence 63, Appl |
| 11 | 181.2 | 10.4 | 2413 | US-11-400-825-1 | Sequence 1, Appl |
| 12 | 178.4 | 10.2 | 2916 | US-10-560-723-98 | Sequence 98, Appl |
| 13 | 178.4 | 10.2 | 3003 | US-10-560-723-96 | Sequence 96, Appl |
| 14 | 178.4 | 10.2 | 3102 | US-10-560-723-97 | Sequence 97, Appl |
| 15 | 178.4 | 10.2 | 3106 | US-10-560-723-95 | Sequence 95, Appl |
| 16 | 178.4 | 10.2 | 3195 | US-11-266-748A-27757 | Sequence 27757, A |
| 17 | 178.4 | 10.2 | 3226 | US-10-560-723-94 | Sequence 94, Appl |
| 18 | 161.4 | 9.2 | 2003 | US-11-266-748A-357278 | Sequence 357278, A |
| 19 | 161.4 | 9.2 | 2003 | US-11-266-748A-440657 | Sequence 440657, A |
| 20 | 154.2 | 8.6 | 2544 | US-11-400-825-3 | Sequence 3, Appl |
| 21 | 150.4 | 8.6 | 2810 | US-11-266-748A-31444 | Sequence 31444, A |
| 22 | 148.4 | 8.5 | 1713 | US-11-266-748A-253384 | Sequence 253384, A |
| 23 | 148.4 | 8.5 | 1713 | US-11-266-748A-276391 | Sequence 276391, A |

| | | | | | | | |
|---|----|-------|-----|------|---|-----------------------|--------------------|
| C | 24 | 148.4 | 8.5 | 1713 | 8 | US-11-266-748A-313901 | Sequence 313901, A |
| | 25 | 144.6 | 8.3 | 2175 | 8 | US-11-293-697-559 | Sequence 559, App |
| | 26 | 143 | 8.2 | 1596 | 6 | US-10-532-668-1 | Sequence 1, Appl |
| | 27 | 135.4 | 7.7 | 1199 | 8 | US-11-266-748A-57061 | Sequence 57061, A |
| | 28 | 135.4 | 7.7 | 1199 | 8 | US-11-266-748A-62066 | Sequence 62066, A |
| | 29 | 134.4 | 7.7 | 639 | 8 | US-11-266-748A-47172 | Sequence 47172, A |
| | 30 | 134.4 | 7.7 | 639 | 8 | US-11-266-748A-209672 | Sequence 209672, A |
| | 31 | 133.4 | 7.6 | 2788 | 6 | US-10-553-436-98 | Sequence 98, Appl |
| | 32 | 128.4 | 7.3 | 786 | 9 | US-11-359-554-1 | Sequence 1, Appl |
| | 33 | 123.6 | 7.1 | 847 | 8 | US-11-266-748A-226483 | Sequence 226483, A |
| | 34 | 123.4 | 7.1 | 856 | 8 | US-11-266-748A-259941 | Sequence 259941, A |
| | 35 | 123.4 | 7.1 | 856 | 8 | US-11-266-748A-320458 | Sequence 320458, A |
| | 36 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-115746 | Sequence 115746, A |
| | 37 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-157910 | Sequence 157910, A |
| | 38 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-224573 | Sequence 224573, A |
| | 39 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-285377 | Sequence 285377, A |
| | 40 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-336806 | Sequence 336806, A |
| | 41 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-395609 | Sequence 395609, A |
| | 42 | 121.6 | 7.0 | 1000 | 8 | US-11-266-748A-466655 | Sequence 466655, A |
| | 43 | 120.8 | 6.9 | 944 | 9 | US-11-359-554-5 | Sequence 5, Appl |
| | 44 | 120.4 | 6.9 | 1122 | 9 | US-11-359-858-1 | Sequence 1, Appl |
| | 45 | 120.4 | 6.9 | 1336 | 8 | US-11-266-748A-85241 | Sequence 85241, A |

ALIGNMENTS

RESULT 1
US-11-266-748A-22522
; Sequence 22522, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Hartin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662, 276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700, 293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22522
; LENGTH: 2393
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-22522
Query Match 95.6%; Score 1670.4; DB 8; Length 2393;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1671, Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CTCGAGACCATGAGAGGACGACCGGAAATGATCTCCAGCAGAACCTTCAGC 60
DB 78 CTCGAGACCATGAGAGGACGACCGGAAATGATCTCCAGCAGAACCTTCAGC 137
OY 61 TGGAGCATCTCCAGCCAGGACATCTCCAGCTGGAGACCTTCAGCGGGGATCTCCAGC 120

OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-112692

Query Match 40.6%; Score 709.8; DB 8; Length 930;
Best Local Similarity 97.5%; Pred. No. 9.8e-155;
Matches 742; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

```
QY 236 CCAGCCAGGATCTCCAGCCCGGGCATCTCCGGCTCTGGGATCATCTTCAGGTCCTCA 295
DB 1 CCAGCCAGGATCTCCAGCCCGGGCATCTCCGGCTCTGGGATCATCTTCAGGTCCTCA 60
QY 296 TCCGCGAGGTCATCATCCGCGAGGTCAAGCTCTGGTGAACAACCTCCCAACGAGGTAC 355
DB 61 TCCGCGAGGTCATCATCCGCGAGGTCAAGCTCTGGTGAACAACCTCCCAACGAGGTAC 120
QY 356 CTGTGTAGAGCAACACCAAGTGGGGCTGTATCCATCCGATCATCTCTGCCAGTACCA 415
DB 121 CTGTGTAGAGCAACACCAAGTGGGGCTGTATCCATCCGATCATCTCTGCCAGTACCA 180
QY 416 CCAGCAACCGAGGCGCACCGAGGAGAGCCAGGTACGAGCTGCCAGTTCACTGGCGG 475
DB 181 CCAGCAACCGAGGCGCACCGAGGAGAGCCAGGTACGAGCTGCCAGTTCACTGGCGG 240
QY 476 GAGGCGCGAAGAGAGCTACCGCTCATCGGCTGCTGCTCTCATTTGCCCTGGTGT 535
DB 241 GAGGCGCGAAGAGAGCTACCGCTCATCGGCTGCTGCTCTCATTTGCCCTGGTGT 300
QY 536 TCGCTCATCATCTCTTTCAGTTCTGGAGGGCCACACAGGATCAGTACAGAGCAG 595
DB 301 TCGCTCATCATCTCTTTCAGTTCTGGAGGGCCACACAGGATCAGTACAGAGCAG 360
QY 596 AGGAGAGCTCTCCCAAGCAGCTCTGCTGTACCGGGGTGGTGAAGTGAAGCTGAG 655
DB 361 AGGAGAGCTCTCCCAAGCAGCTCTGCTGTACCGGGGTGGTGAAGTGAAGCTGAG 420
QY 656 AGTACAGAGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 715
DB 421 AGTACAGAGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480
QY 716 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAG 775
DB 481 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAG 540
QY 776 AAGACTCTGCGCAGCTGGGTTTGGAGGTCTCAACCGGAGAACCGAGTTGCCACAGG 835
DB 541 AAGACTCTGCGCAGCTGGGTTTGGAGGTCTCAACCGGAGAACCGAGTTGCCACAGG 600
QY 836 GATTTTGCACAGCTTCTCATCTTGAGATACAACTCCACATCCAGAGAGCTCCAC 895
DB 601 GATTTTGCACAGCTTCTCATCTTGAGATACAACTCCACATCCAGAGAGCTCCAC 660
QY 896 AGGTCTGAGTCCCTTCCAGCGGTATATCT-CCCTCCAGTGTCCACT-GCGAGCTGA 953
DB 661 AGGTCTGAGTCCCTTCCAGCGGTATATCTCCCTCCAGTGTCCACTGCGGAGCTGA 720
QY 954 GGGCCATGACCGGCGGATCTGGAGGGGCGCTGGCTCG 994
DB 721 GGGCCATGACCGGCGGATCTGGAGGGGCGCTGG 761
```

RESULT 3

US-11-266-748A-11696/c
Sequence 11696, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnson, Patrick
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 11696
LENGTH: 255
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-11696

Query Match 13.7%; Score 240.2; DB 8; Length 255;
Best Local Similarity 98.8%; Pred. No. 3.4e-46;
Matches 242; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 514 CTCTCTATGCTCCCTGGTGTGTTGCTTCATCATCTCTTCCAGTTCTGGCAGGCGCACAC 573
DB 255 CTCTCTATGCTCCCTGGTGTGTTGCTTCATCATCTCTTCCAGTTCTGGCAGGCGCACAC 196
QY 574 AGGATAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
DB 195 AGGATAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 136
QY 634 GGTGTGAGCTGACAGCTGAAGAGTGAAGAGCTGGCTGGTGGTGGTGGTGGTGGTGGT 693
DB 135 GGTGTGAGCTGACAGCTGAAGAGTGAAGAGCTGGCTGGTGGTGGTGGTGGTGGTGGT 76
QY 694 GTCTCTCTTAAATCTACTCTGGGCTCTCCATCAGTGGCTTCCATCTGTAGACGAA 753
DB 75 GTCTCTCTTAAATCTACTCTGGGCTCTCCATCAGTGGCTTCCATCTGTAGACGAA 16
QY 754 CTGGA 758
DB 15 CAGCA 11
```

RESULT 4

US-11-266-748A-64397
Sequence 64397, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnson, Patrick
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2

```

1  PRIOR FILING DATE: 2004-11-03
2  PRIOR APPLICATION NUMBER: US 60/662,276
3  PRIOR FILING DATE: 2005-03-14
4  PRIOR APPLICATION NUMBER: US 60/700,299
5  PRIOR FILING DATE: 2005-07-18
6  NUMBER OF SEQ ID NOS: 48396
7  SOFTWARE: PatentIn version 3.3
8  SEQ ID NO 64397
9  LENGTH: 255
10  TYPE: DNA
11  ORGANISM: Homo Sapiens
12  US-11-266-748A-64397

```

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 13.7% | Score 240.2; | DB 8; | Length 255; |
| Best Local Similarity | 98.8% | Pred. No. 3.4e-46; | | |
| Matches 242; Conservative | 0; | Mismatches 3; | Indels 0; | Gaps 0; |

| QY | 514 | CGCTCTCATTGACCTCTGATGATTTTCGCTCATCATCTCTTCCAGTCTGAGAGGGCACAC | 574 |
|----|-----|---|-----|
| Db | 1 | CTCTCATATGCCCTCGTGGTTTCGCTCATCATCTCTTCCAGTCTGAGAGGGCACAC | 60 |
| QY | 574 | AGGGAATCAGGTTACAAAGAGCACAAGGAGAGCGTCTCCCAAGCACGCTGTTCCCTGTGACCG | 633 |
| Db | 61 | AGGGAATCAGGTTACAAAGAGCACAAGGAGAGCGTCTCCCAAGCACGCTGTTCCCTGTGACCG | 120 |
| QY | 634 | GGTGTGTGACTGTCAAGCTGTAAGAGTGCACAGCTGGGCTGCGTGAAGTTTGACTGGGACAA | 693 |
| Db | 121 | GGTGTGTGACTGTCAAGCTGTAAGAGTGCACAGCTGGGCTGCGTGAAGTTTGACTGGGACAA | 180 |
| QY | 694 | GTCTCTGCTTTAAATCTACTCTTGGGCTCTCCCATCAGTAGTGGCTTCCCATCTGTAGACGAA | 753 |
| Db | 181 | GTCTCTGCTTTAAATCTACTCTTGGGCTCTCCCATCAGTAGTGGCTTCCCATCTGTAGACGAC | 240 |
| QY | 754 | CTGGA | 758 |
| Db | 241 | CAGCA | 245 |

```

1 RESULT 5
2 US-11-266-748A-67229/C
3 Sequence 67229, Application US/11266748A
4 Publication NO. US20060134663A1
5 GENERAL INFORMATION:
6 APPLICANT: Hartin, Paul
7 APPLICANT: Johnston, Patrick
8 APPLICANT: Mulligan, Karl
9 TITLE OF INVENTION: Transcryption Microarray Technology and
10 FILE REFERENCE: 55815-0102 (319189)
11 CURRENT APPLICATION NUMBER: US/11/266,748A
12 CURRENT FILING DATE: 2005-11-03
13 PRIOR APPLICATION NUMBER: EP 04105479.2
14 PRIOR FILING DATE: 2004-11-03
15 PRIOR APPLICATION NUMBER: EP 04105482.6
16 PRIOR FILING DATE: 2004-11-03
17 PRIOR APPLICATION NUMBER: EP 04105483.4
18 PRIOR FILING DATE: 2004-11-03
19 PRIOR APPLICATION NUMBER: EP 04105507.0
20 PRIOR FILING DATE: 2004-11-03
21 PRIOR APPLICATION NUMBER: EP 04105485.9
22 PRIOR FILING DATE: 2004-11-03
23 PRIOR APPLICATION NUMBER: EP 04105484.2
24 PRIOR FILING DATE: 2004-11-03
25 PRIOR APPLICATION NUMBER: US 60/662,216
26 PRIOR FILING DATE: 2005-03-14
27 PRIOR APPLICATION NUMBER: US 60/700,293
28 PRIOR FILING DATE: 2005-07-18
29 NUMBER OF SEQ ID NOS: 483996
30 SOFTWARE: PatentIn version 3.3
31 SEQ ID NO 67229
32 LENGTH: 255
33 TYPE: DNA
34 ORGANISM: Homo Sapiens

```

US-11-266-748A-67229

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 13.7% | Score 240.2; | DB 8; | Length 255; |
| Best Local Similarity | 98.8% | Pred. No. 3.4e-46; | | |
| Matches 242; Conservative | 0; | Mismatches 3; | Indels 0; | Gaps 0 |

| | | | |
|----|-----|---|-----|
| Qy | 511 | CCCTCCCATATGGCCCTGGTGGTTCCGTCATCATCTCTTCCAGTTCTGGCAGGGCCAC | 578 |
| Db | 255 | CCCTCCCATATGGCCCTGGTGGTTTGGCTATCATCTCTTCCAGTTCTGGCAGGGCCAC | 196 |
| Qy | 574 | AGGGATCAGGTACAGGAGCAGAGGAGAGCTGTCTCCACGACGCTGTTCCGCTGTGACGG | 633 |
| Db | 195 | AGGGATCAGGTACAGGAGCAGAGGAGAGCTGTCTCCACGACGCTGTTCCGCTGTGACGG | 136 |
| Qy | 634 | GGTGTGTGACCTGCAAGCTGAAGAGTGAAGACGTGGGCTGCGTGAAGTTTGACTGGGACAA | 693 |
| Db | 135 | GGTGTGTGACCTGCAAGCTGAAGAGTGAAGACGTGGGCTGCGTGAAGTTTGACTGGGACAA | 76 |
| Qy | 694 | GTTCTTGCTTTAAATTTACTCTGGGGTCTCTCCCATATAGTGGCTTCCCATCTGTACAGCAA | 753 |
| Db | 75 | GTTCTTGCTTTAAATTTACTCTGGGGTCTCTCCCATATAGTGGCTTCCCATCTGTACAGCAAC | 16 |
| Qy | 754 | CTGGA 758 | |
| Db | 15 | CAGCA 11 | |

```

RESULT 6
US-10-196-749-329
Sequence 329, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Deenoyers, Inc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 329
LENGTH: 2063
TYPE: DNA

```

ORGANISM: Homo Sapien
US-10-196-749-329

Query Match 11.1%; Score 193.2; DB 6; Length 2063;
Best Local Similarity 55.8%; Pred. No. 5.5e-35;
Matches 435; Conservative 0; Mismatches 333; Indels 12; Gaps 3;

```
QY 924 TCTCCCTCCAGTGTCCCACTGGAGCGGACCTGAGCGGCGGAGTCCGAGGAGG 983
DB 760 TCTCCCTCCAGTGTCCCACTGGAGCGGACCTGAGCGGCGGAGTCCGAGGAGG 819
QY 984 CGCTGGCTCCGAGATGACAGAGTGGCTTGGCAAGTGTGCTGCTGCTGCGACCA 1043
DB 820 AGAGAGCTCTGTGATTTCTTGCTTGGCAGGTAGAGTCAAGTACGAAACAGCAG 879
QY 1044 TCTGTGAGGACGCTCATTTACCGCCAGTGGTGTCTGCTGCTGCGCCACTGCTTTCG 1103
DB 880 TCTGTGAGGAGGACCTCTGAGCCCGCACTGGGTCTTCAAGCGGACCGCACTGCTTC--- 935
QY 1104 TGACCCGGGAGAGAGTCTCTGAGGCGTGGAGGTGTAGCGGGGACCGAACCTGCAAC 1163
DB 936 --AGAGAAACATACCGATGTGTTCACATGGAAGTGGCGGAGCTCAGACAAACTGGGCA 993
QY 1164 AGTGGCTGAGGACGCTCCATTCGCGAGATCATCATCAAGCAAGCAATTAACCCGATGAG 1223
DB 994 GCTTCCC---ATCCCTGGCTGTGGCGCAAGATCATCATTAATTAACCCGATGAG 1050
QY 1224 AGAGCACTATGACATGCGCTCATGCGGCTGTCCAGCCCTTGAACCTGTCCGCTACA 1283
DB 1051 CCAAGACATATGACATGCGCTCATGAGCTGAGCTGCACTCACTTTCAGGACAG 1110
QY 1284 TCCACCTGCTTGTCTCCCATGATGACAGACCTTTAGCTCAATAGACCTGTCTGGA 1343
DB 1111 TCAGGCCATGTGTCTGCTTCTTTTATGAGAGACTCACTCCAGCCACCCCACTTGA 1170
QY 1344 TCACAGGCTTTGGCAAGACAGAGGAGACAGATGACAAACATCCCTTCCGCGAGG 1403
DB 1171 TCATTTGATGGGGCTTTTACAGAGCAAGATGAGAGGAATGTCTGACATATCTGTGACG 1230
QY 1404 TGCAGGTCAATCTCATGACTTCAAGAAATGCAATGACTTGTGTATGACAGTTACC 1463
DB 1231 CGTCAGTCCAGATGATGACAGACACAGGTGCAATGACAGATGCGTACAGGGGAG 1290
QY 1464 TTATCCCAAGATGATGTGTGTGTGGGAGCTTTCGTGGGGGAGAGACTTCTCCAGGAG 1523
DB 1291 TCACCGAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1350
QY 1524 ACAGCGGGGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1583
DB 1351 ACAAGTGTGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1407
QY 1584 GCTGGGGGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1643
DB 1408 GCTGGGGGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1467
QY 1644 TTCTTCTCTGATTTTACGCAAGATGAGAGGAGGTGCTGATTCAGAAATCTTACAG 1703
DB 1468 ATCTCACTGATCTTCAATGTCTGTGAAGGCTGAGCTGTATGTGTGTGTGTGTGTGT 1527
```

RESULT 7

US-11-101-316-111
Sequence 111, Application US/11101316
Publication No. US2006009657A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Guernsey, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
FILE REFERENCE: P3330R1C17C1

CURRENT APPLICATION NUMBER: US/11/101,316
CURRENT FILING DATE: 2005-04-06
PRIOR APPLICATION NUMBER: 10/063526
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 111
LENGTH: 2063
TYPE: DNA
ORGANISM: Homo Sapien
US-11-101-316-111

Query Match 11.1%; Score 193.2; DB 8; Length 2063;
Best Local Similarity 55.8%; Pred. No. 5.5e-35;
Matches 435; Conservative 0; Mismatches 333; Indels 12; Gaps 3;

```
QY 924 TCTCCCTCCAGTGTCCCACTGGAGCGGACCTGAGCGGCGGAGTCCGAGGAGG 983
DB 760 TCTCCCTCCAGTGTCCCACTGGAGCGGACCTGAGCGGCGGAGTCCGAGGAGG 819
QY 984 CGCTGGCTCCGAGATGACAGAGTGGCTTGGCAAGTGTGCTGCTGCTGCGACCA 1043
DB 820 AGAGAGCTCTGTGATTTCTTGCTTGGCAGGTAGAGTCAAGTACGAAACAGCAG 879
QY 1044 TCTGTGAGGACGCTCATTTTACCGCCAGTGGTGTCTGCTGCTGCGCCACTGCTTTCG 1103
DB 880 TCTGTGAGGAGGACCTCTGAGCCCGCACTGGGTCTTCAAGCGGACCGCACTGCTTC--- 935
QY 1104 TGACCCGGGAGAGAGTCTCTGAGGCGTGGAGGTGTAGCGGGGACCGAACCTGCAAC 1163
DB 936 --AGAGAAACATACCGATGTGTTCACATGGAAGTGGCGGAGCTCAGACAAACTGGGCA 993
QY 1164 AGTGGCTGAGGACGCTCCATTCGCGAGATCATCATCAAGCAAGCAATTAACCCGATGAG 1223
DB 994 GCTTCCC---ATCCCTGGCTGTGGCGCAAGATCATCATTAATTAACCCGATGAG 1050
QY 1224 AGAGCACTATGACATGCGCTCATGCGGCTGTCCAGCCCTTGAACCTGTCCGCTACA 1283
DB 1051 CCAAGACATATGACATGCGCTCATGAGCTGAGCTGCACTTCCCACTCACTTTCAGGACAG 1110
QY 1284 TCCACCTGCTTGTCTCCCATGATGACAGACCTTTAGCTCAATAGAGACCTGTCTGGA 1343
DB 1111 TCAGGCCATGTGTCTGCTTCTTTTATGAGAGACTCACTCCAGCCACCCCACTTGA 1170
QY 1344 TCACAGGCTTTGGCAAGACAGAGGAGACAGATGACAAACATCCCTTCCGCGAGG 1403
DB 1171 TCATTTGATGGGGCTTTTACAGAGCAAGATGAGAGGAATGTCTGACATATCTGTGACG 1230
QY 1404 TGCAGGTCAATCTCATGACTTCAAGAAATGCAATGACTTGTGTATGACAGTTACC 1463
DB 1231 CGTCAGTCCAGATGATGACAGACACAGGTGCAATGACAGATGCGTACAGGGGAG 1290
QY 1464 TTATCCCAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1523
DB 1291 TCACCGAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1350
QY 1524 ACAGCGGGGGCTCTTGTCTGTGAGCAAGAAACAGCTGTGTATCTGCGAGGTGTACCA 1583
DB 1351 ACAAGTGTGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1407
QY 1584 GCTGGGGGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1643
DB 1408 GCTGGGGGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1467
```


03-10-196-149-63

NUMBER OF SEQ ID NOS: 134

```
/ SEQ ID NO 1
/ LENGTH: 2413
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ OTHER INFORMATION: entire cDNA sequence of TADG-12 gene
US-11-400-825-1
```

```
Query Match      10.4%; Score 181.2; DB 7; Length 2413;
Best Local Similarity 55.6%; Pred. No. 3.5e-32;
Matches 414; Conservative 0; Mismatches 318; Indels 12; Gaps 3;
```

```
QY 953 AGGGCCATGACCGGGCGGATGCTGGGAGAGGGCGCTGGCTCGGATGACAGTGGCTTGG 1012
Db 777 AGGGGCTACAGCTCAGCATCGATGGTGGTGAAGAAATGTCCTTGTCTGCGACAGTGGCCCTGG 836
QY 1013 CAATGAGTCTGCACTTGGGACCAACCAATCTGTGGAGGACGCTCATTTGACGCCAG 1072
Db 837 CAGGCGACCTTCAAGTTCAGAGGCTACCACTGTGCGGGGCTCTGTATCATCACGCCCTTG 896
QY 1073 TGGGAGTCACTGCGCCGACATGCTTCTTGTGACCCGGAGAAAGTCTGAGGGCTGG 1132
Db 897 TGAATCATCTGCTGCACTGCTGTATG-----ACTTGTACTTCCCAAGTCATGG 950
QY 1133 AAGGTGTACGGGGGACCGACCAACCTTGACCAAGTTGCTTGAAGGACGCTTCATTTGCCAG 1192
Db 951 ACCATTCAGGTGGGTCTAGTTTCCCTGTGTGACATCCAGGCCCATTCCTCATTTGGTGGAG 1010
QY 1193 ATCA--TCATCAACAGCAATTTACACCGATAGAGAGACATATGATCATGCTCCCTCATG 1249
Db 1011 AAGATTGTCTACACAGCAAGTATACAGCAAGAGGCTGGGCAATGATCATGCTCCCTTATG 1070
QY 1250 CGGCTGTCCAAAGCCCTGTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
Db 1071 AAGCTGGCCGGGACCACTACGTTCAATGAATGATCCAGCTGTGTGCTGCTGCTGCTGCTGCT 1130
QY 1310 GGACAGACCTTTAAGCTTCATATGAGACCTGTGATCAGAGGCTTTGGCAAGACAGAGGAG 1369
Db 1131 GAAAGAACTTCCCGATGAGAAAGTGTGTGAGACGTAGAGATGGGGGGCCACAGAGGAT 1190
QY 1370 ACAGATGACAAAGACATCCCTCTCTCCGGAGGTGACGTCAATCTCATGACTTCAAG 1429
Db 1191 GGAGGTGAC---GCTCTCCCTGTCTGTCAACACAGCGGCGCTCCCTTGTATTTCCAAAG 1247
QY 1430 AATGCAATGACTACTTGTCTATGACAGTTACCTTACCCCAAGAGATGATGTGTGGG 1489
Db 1248 ATCTGCAACCAAGAGGAGCTGTACGCTGACATCTCTCCCTCAATGCTCTGCGCGGGC 1307
QY 1490 GACCTTCTGGGGGAGAGACTCTCTGCGAGGAGAGACAGCGGGGGCTCTTGTCTGTGAG 1549
Db 1308 TACCTGACCGGTGTGGTAACAGCTGCGAGGGGAGACAGCGGGGGCTCTGTGTGTCAA 1367
QY 1550 CAGAACACCGCTGTACTCTGGCAGGTGTACACAGCTGGGGCAACAGGCTGTGGCCAGAGA 1609
Db 1368 GAGAGGAGGCTGTGGAAGTTAGTGGAGCGACACGCTTGTGCATCTGCTGCGCAGAGGTG 1427
QY 1610 AACAAACCTGTGTGTATACCAAGTGTACAGAGTTCTTCCCTGATTTACAGCAAGTGT 1669
Db 1428 AACAAAGCTGGGGGTGTACACCGGTGTACCTTCTTCTGATGATCAGAGCAAGTGT 1487
QY 1670 GAGAGCGAGGTGCAATTCAGAAAA 1693
Db 1488 GAGAGAGACCTTAAAAAAGCTGAAAG 1511
```

```
RESULT 12
US-10-560-723-98
; Sequence 98, Application US/10560723
; Publication No. US20060115821A1
; GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
```

```
/ TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
/ FILE REFERENCE: 80213WO
/ CURRENT APPLICATION NUMBER: US/10/560,723
/ CURRENT FILING DATE: 2005-12-15
/ NUMBER OF SEQ ID NOS: 185
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 98
/ LENGTH: 2916
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-560-723-98
```

```
Query Match      10.2%; Score 178.4; DB 6; Length 2916;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;
```

```
QY 969 GGATCGTGGAGGGGCGCTGCTCGATAGCAAGTGGCTTGGCAAGTGTGTGCACT 1028
Db 581 GGATCGTGGGGGTGAAGAGCGCGCTCCGGGGGCTGGCCCTGGCAGGTGAGCTTGACAG 640
QY 1029 TCGGACACCCACATCTGTGTGAGAGCAAGCTATTGACGCCCAATGGGTGCTCACTGCG 1088
Db 641 TCCGAAGCTTCACGCTGTGCGAGGCTTCATCATCACCCGAGTGGATCGTACAGCG 700
QY 1089 CCACTGCTTCTTGTGACCCGGAGAGAGTCTGAGAGGCTGGAAGGTGTACCGGGGCA 1148
Db 701 CCACTGCTGTGAAAAAAGCTTTAAATCATGATGGAAGGCAATTGGACGGGGAATT 760
QY 1149 CCAGCAACCTGACCAAGTTGCTTGAAGCAGCTTCATTTGCCAGAA---TCATCATCAACA 1205
Db 761 TGAGACATCTTTATGTATGTATGAGCCGAGTACCAAGTGAAGAAAGTGTATCTATC 820
QY 1206 GCAATTCACCCGATGAGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAAGCCC 1265
Db 821 CAATTTATGACTCCAAAGACCAAGAACATGACATTTGGCTGTATGAAGTGCAGAAAGCTC 880
QY 1266 TGACCTGTGCGCTACATCCACCTGCTTGTCTCCCATGATGACAGACCTTTAGCC 1335
Db 881 TGACTTTCAAGCACTGTGTGAAACAGTGTGTGTGCGCAACCCAGGCAATGATGTGCGAGC 940
QY 1326 TCAATGAGACCTGTGTGATCAGAGCTTTGGCAAGACCAAGGAGACAGATGTACAAGACAT 1385
Db 941 CAGAACAGCTCTGTGTGATTTCCGGGTGGGGGACACCGAGAGAA---AAGGGAAGACT 997
QY 1386 CCCCCTTCTCCGGGAGGTGAGCTCATCTCATGCACTTCAAGAAATGCAATGACTACT 1445
Db 998 CAGAAAGTCTGAACAGCTGCGCAAGTGTCTTCTCATTTGAGACACAGAGATGCAACAGCAT 1057
QY 1446 TGATCTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGACCTTGTGTGGGGCA 1505
Db 1058 ATGTCTATGACCACTGATCAACACAGCATGATCTGTGCGGCTTCTGTGAGGGGAAAG 1117
QY 1506 GAGACTCTGCGCAGAGGAGACAGCGGGGGCTCTTGTCTGTGAGCAGAAACACCGCTGTGT 1565
Db 1118 TCGATTCTTTCAGAGGAGAGAGTGAAGGAGGCTCTGTGTGATCTTCAAGAAACAAATCTGTGT 1177
QY 1566 ACCCTGGAGGTGTACCAAGCTGGGGGCAAGGCTGTGGCAGAGAAACAACTGTGTGT 1625
Db 1178 GGTGTATAGGGGATACAAAGCTGGGGTTCTGTCTGTGCAAGGCTTACAGACAGAGATGT 1237
QY 1626 ACACCAAGTGTACGAAGTCTTCTGATTTTACAGCAAGATGAGAGCGAGGTGCGAT 1685
Db 1238 ACAGGAAATGTGATGTATTCAGCGGACTGTGATTTTACAAATGAGAGGAGAGACGGCTAAT 1287
QY 1686 TCAGAAATCTT 1697
Db 1298 CCACATGTGCTT 1309
```

```
RESULT 13
US-10-560-723-96
; Sequence 96, Application US/10560723
; Publication No. US20060115821A1
```


GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: B0213WO
; CURRENT APPLICATION NUMBER: US/10/560,723
; CURRENT FILING DATE: 2005-12-15
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-560-723-96

Query Match 10.2%; Score 178.4; DB 6; Length 3003;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;

QY 969 GGATCGTGGAGGGGCGCTGCGCTCGATGACAGTGGCCCTTGGCAAGTGAAGTGCACCT 1028
DB 668 GGATCGTGGAGGGGCGCTGCGCTCGATGACAGTGGCCCTTGGCAAGTGAAGTGCACCG 727
QY 1029 TCGGACACCAACCATCTGTGTGAGAGCAAGCTATTGACGCCAGTGGGTCTCACTGCGG 1088
DB 728 TCCGAAAGCTTCAAGTGTGCGAGGCTTCATCATCACCCTCGAGTGAATCTGACAGCCG 787
QY 1089 CCCACTGCTTCTTGCTGTAACCGGAGAAAGTCTGAGAGGCTGGAAGGTATACGCGGCA 1148
DB 788 CCCACTGCTTGGAAAAACCTTTTAACAATTCATGGCATTTGACGGCATTTGCGGGGATTT 847
QY 1149 CCAGCAACCTGACCAAGTTCCTTGAAGGAGCTTCATTGCCGAGA--TCAATCATCAACA 1205
DB 848 TGAACAATCTTTTATGTTTATGAGAGCGGATACCAAGTAAAAAGTGAATTTCTCATTC 907
QY 1206 GCAATTCACCGATGAGAGGAGACGATATGATCGCCCTCATGCGGCTGTCCAGCCCC 1265
DB 908 CAAATTTATGATCCAAAGCAAGCAATGACATTTGCGCTATGAGTGCAGAGGCTTC 967
QY 1266 TGACCTGTCCGCTCAACATCCACCTGCTGCTCCCTCCATGATGAGACACTTTAGCC 1325
DB 968 TGACTTTTCAAGCACTTATGAAACCAAGTGTCTGCCCAACCGAGCATGATGTGCAGC 1027
QY 1326 TCAATGAGACTGCTGTGATCAAGGCTTTGGCAAGACCAAGGAGACAGATGACAGACAT 1385
DB 1028 CAGAACAGCTCTGCTGGAATTTCCGGGTGGGGGCCACGAGAGAG--AAGGGAAGACT 1084
QY 1386 CCCCCCTTCGCGGAGGTGAGGATCATCTCATGACTTCAAGAAATGCAATGACTACT 1445
DB 1085 CAGAAAGTCTGAACGCTGCCAAGGTGCTTCTCATTTGAGACACAGAGATGCAACAGCAT 1144
QY 1446 TGCTCTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGACCTTGTGGGGCA 1505
DB 1145 ATGTCTATGACAACTGATCAACACAGCATGATGTGTGCGGCTTCTGCAAGGGAGACG 1204
QY 1506 GAGACTCTGCGCAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGCAAAACCGCTGGT 1565
DB 1205 TCGATTTCTTGGCAGGGGTACAGTGGAGGCGCTGTGCTCATTCGAAAGAACATATCTGGT 1264
QY 1566 ACCTGGCAGGTGTCAACGAGCTGGGGCAAGGCTGTGGCAGAGAAACAACTGGTGTGT 1625
DB 1265 GCGTGTATGAGGAGATACAGCTGGGGTCTGGCTGTGCCAAGACTTACAGACCCAGAGGT 1324
QY 1626 ACACCAAGTGAAGAGTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGCAT 1685
DB 1325 ACGGGAATGTATGTATTTCAACGAGCTGATTTATCGAACAAATAGAGGAGACGCGCTAAT 1384
QY 1686 TCAGAAAATCTCT 1697
DB 1385 CCACATGTGCTTT 1396

RESULT 14

US-10-560-723-97
; Sequence 97; Application US/10560723
; Publication No. US20060115821A1
; GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: B0213WO
; CURRENT APPLICATION NUMBER: US/10/560,723
; CURRENT FILING DATE: 2005-12-15
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 3102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-560-723-97

Query Match 10.2%; Score 178.4; DB 6; Length 3102;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;

QY 969 GGATCGTGGAGGGGCGCTGCGCTCGATGACAAAGTGGCCCTTGGCAAGTGAAGTGCACCT 1028
DB 767 GGATCGTGGAGGGGCGCTGCGCTCGATGACAAAGTGGCCCTTGGCAAGTGAAGTGCACCG 826
QY 1029 TCGGACACCAACCATCTGTGTGAGAGCAAGCTATTGACGCCAGTGGGTCTCACTGCGG 1088
DB 827 TCCGAAAGCTTCAAGTGTGCGAGGCTTCATCATCACCCTCGAGTGAATCTGACAGCCG 886
QY 1089 CCCACTGCTTCTTGCTGTAACCGGAGAAAGTCTGAGAGGCTGGAAGGTATACGCGGCA 1148
DB 887 CCCACTGCTTGGAAAAACCTTTTAACAATTCATGGCATTTGACGGCATTTGCGGGGATTT 946
QY 1149 CCAGCAACCTGACCAAGTTCCTTGAAGGAGCTTCATTGCCGAGA--TCAATCATCAACA 1205
DB 947 TGAACAATCTTTTATGTTTATGAGAGCGGATACCAAGTAAAAAGTGAATTTCTCATTC 1006
QY 1206 GCAATTCACCGATGAGAGGAGACATATGATCGCCCTCATGCGGCTGTCCAGCCCC 1265
DB 1007 CAAATTTATGATCCCAAGCAAGCAATGACATTTGCGCTATGAGTGCAGAGGCTTC 1066
QY 1266 TGACCTGTCCGCTCAACATCCACCTGCTGCTCCCTCCATGATGAGACACTTTAGCC 1325
DB 1067 TGACTTTTCAAGCACTTATGAAACCAAGTGTCTGCCCAACCGAGCATGATGTGCAGC 1126
QY 1326 TCAATGAGACTGCTGTGATCAAGGCTTTGGCAAGACCAAGGAGACAGATGACAGACAT 1385
DB 1127 CAGAACAGCTCTGCTGGAATTTCCGGGTGGGGGCCACGAGAGAG--AAGGGAAGACT 1183
QY 1386 CCCCCCTTCGCGGAGGTGAGGATCATCTCATGACTTCAAGAAATGCAATGACTACT 1445
DB 1184 CAGAAAGTCTGAACGCTGCCAAGGTGCTTCTCATTTGAGACACAGAGATGCAACAGCAT 1243
QY 1446 TGCTCTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGACCTTGTGGGGCA 1505
DB 1244 ATGTCTATGACAACTGATCAACACAGCATGATGTGTGCGGCTTCTGCAAGGGAGACG 1303
QY 1506 GAGACTCTGCGCAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGCAAAACCGCTGGT 1565
DB 1304 TCGATTTCTTGGCAGGGGTACAGTGGAGGCGCTGTGCTCATTCGAAAGAACATATCTGGT 1363
QY 1566 ACCTGGCAGGTGTCAACGAGCTGGGGCAGGCTGTGGCAGAGAAACAACTGGTGTGT 1625
DB 1364 GCGTGTATGAGGAGATACAGCTGGGGTCTGGCTGTGCCAAGCTTACAGACCCAGAGGT 1423
QY 1626 ACACCAAGTGAAGAGTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGCAT 1685
DB 1424 ACGGGAATGTATGTATTTCAACGAGCTGATTTATCGAACAAATAGAGGAGACGCGCTAAT 1483
QY 1686 TCAGAAAATCTCT 1697
DB 1484 CCACATGTGCTTT 1495

RESULT 15
US-10-560-723-95
; Sequence 95: Application US/10560723
; Publication No. US20060115821A1
; GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: B0213M0
; CURRENT APPLICATION NUMBER: US/10/560,723
; CURRENT FILING DATE: 2005-12-15
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-560-723-95

Query Match 10.2%; Score 178.4; DB 6; Length 3106;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;

QY 969 GGATCGTGGAGGAGGCGCTGGCTCGATAGCAAGTGGCTTTGGCAAGTGAAGTCTGCACT 1028
DB 771 GGATCGTGGGCGGTGAGAGCGCGCTCCCGGGGCTGGCCCTGGCAGGTCAAGCTGCACG 830
QY 1029 TCGGCACACCCACATCTGTGAGAGCAAGCTCATTTAGCGCCAGTGGGTGCTCACTGCCG 1088
DB 831 TCCAGAAAGTCCACGTGTGGAGGAGCTTCATCATCACCCCGAGTGAATGTTGACAGCCG 890
QY 1089 CCCACTGCTTCTTGTGTGACCCGAGGAAAGTCTGSAAGGCTGGAAGTGTACCGGGCA 1148
DB 891 CCCACTGCGTGGAAAAACCTCTTAACAATCATGCAATTGGAACGCAATTTGCGGGAAATT 950
QY 1149 CCAGCAACTGACCAAGTGTGCTGAGGAGCGCTTCATTTCCGAGA---TCATCATCAACA 1205
DB 951 TGAGACATATCTTTCATGTCTATGAGCCGATACCAAGTAGAAAAAGTATTTCTCATC 1010
QY 1206 GCATTTACCCGATGAGAGAGACATATGATTCGCGCTCATCGGCTGTCCAAAGCCC 1265
DB 1011 CAATTTATGATCTCAAGACCAAGAACATATGATTCGCTGATGAGCTGCAAGGCTTC 1070
QY 1266 TGAACCTGTCCGCTCAATCCACCTGCTTGCCTCCCATGATGACAGACTTTAGCC 1325
DB 1071 TGAATTTCAAGCACTTGTAGTGAACAGTGTGTCTGCCCAACCGCATGTGTGTCAGC 1130
QY 1326 TCAATGAGACTGTCTGTATCAAGGCTTTGGCAAGACCGAGAGACAGATGACAAAGCAT 1385
DB 1131 CAGAACAGCTCTGCTGATTTCCGGGTGGGGGCGCACCGAGAGAG---AAGGGAAGACT 1187
QY 1386 CCCCCCTTCCTCGGAGAGGTGACATCTCATGCACTTCAAGAAATGCAATGACTACT 1445
DB 1188 CAGAAAGTGTGAACGCTGCCAAGGTGCTTCTCATTTGAGACACAGAGATGTCAACAGCAT 1247
QY 1446 TGGTCTATGACAGTTACTTACCCCAAGATGATGTGTGCTGGGAGCTTGTGTGGGCA 1505
DB 1248 ATGTCTATGACCACTGATCAACACAGCATGATCTGTGCCGCTTCTTGCAGGGGAAAG 1307
QY 1506 GAGACTCTGTCAGAGGAGACAGCGGGGGGCTTGTGTGTGAGCAAGAAACCGCTGGT 1565
DB 1308 TCGATTTCTTGGCAAGGTGACAGTGAAGGCTCTGTGTCACTTCGAGAAACAATATCTGGT 1367
QY 1566 ACCGTGGAGGTGTACCAAGCTGGGGGCAAGGCTGTGGCCAGAGAAACAACTGTGTGT 1625
DB 1368 GCGTGTAGGGGATTAAGCTGGGCTTCTGTGTGTGCAAGCTTTACAGACAGAGAGTGT 1427
QY 1626 ACACCAAGTGTACAGAAATTTCTTCCCTGATTTTACAGCAAGATGAGAGCGAGGTGCGAT 1685
DB 1428 ACGGGAATGTGATGTATTCACGGAATGATTTTATTCGACAAATGAGGGCAGACGGCTAAT 1487

QY 1686 TCAGAAAAATCCT 1697
DB 1488 CCACATGTGCTTT 1499

Search completed: September 18, 2006, 13:15:32
Job time : 1952 Secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 01:23:32 ; Search time 58.6 Seconds
(without alignments)
4091.541 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 3333

Sequence: 1 ctccagcaccatggagaggg.....gctcgtcgtgactcgcgagaa 1748

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 5179358

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=x1p
-O=/abs/ABSSWEB.spool/US10806370/runat_15092006_105622_9871/app_query.fasta.1
-DB=A_Geneseq -OFT=fasta -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10806370.CCGN_1_1_440 @runat_15092006_105622_9871 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_8: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *
9: geneseqp2005s: *
10: geneseqp2006s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 3004 | 90.1 | 569 | 8 | ADH17440 Human NOV |
| 2 | 2999 | 90.0 | 562 | 5 | AAE17238 Human tra |
| 3 | 2991 | 89.7 | 562 | 4 | AAE01943 Human tra |
| 4 | 2991 | 89.7 | 562 | 7 | AD110393 Human cel |
| 5 | 2991 | 89.7 | 562 | 8 | ADJ46917 Human tra |
| 6 | 2991 | 89.7 | 562 | 9 | ADY50146 Human end |
| 7 | 2963 | 88.9 | 688 | 4 | AAE01944 Human tra |
| 8 | 2963 | 88.9 | 688 | 7 | AD110395 Human cel |

| | | | | | | |
|----|--------|------|-----|---|----------|--------------------|
| 9 | 2963 | 88.9 | 688 | 8 | ADJ46919 | ADJ46919 Human tra |
| 10 | 2963 | 88.9 | 688 | 9 | ADY50054 | ADY50054 Human end |
| 11 | 2956 | 88.7 | 581 | 7 | ADE31743 | ADe31743 Human 291 |
| 12 | 2953.5 | 88.6 | 593 | 8 | ADH17430 | Adh17430 Human NOV |
| 13 | 2948.5 | 88.5 | 586 | 8 | ADH17414 | Adh17414 Human NOV |
| 14 | 2948.5 | 88.5 | 586 | 8 | ADH17434 | Adh17434 Human NOV |
| 15 | 2943.5 | 88.3 | 586 | 8 | ADH17452 | Adh17452 Human NOV |
| 16 | 2942.5 | 88.3 | 586 | 8 | ADH17450 | Adh17450 Human NOV |
| 17 | 2942 | 88.3 | 581 | 9 | ADZ75552 | ADz75552 Human mos |
| 18 | 2775.5 | 83.3 | 537 | 5 | AAU82746 | AAu82746 Amrno aci |
| 19 | 2775.5 | 83.3 | 537 | 8 | ADH17438 | Adh17438 Human NOV |
| 20 | 2737.5 | 82.1 | 556 | 9 | AEA20240 | Aea20240 Novel hum |
| 21 | 2691 | 80.7 | 542 | 8 | ADH17432 | Adh17432 Human NOV |
| 22 | 2414.5 | 72.4 | 471 | 8 | ABM83371 | ABm83371 Human dia |
| 23 | 2291.5 | 68.8 | 477 | 6 | ABP56842 | ABp56842 Human den |
| 24 | 2287 | 68.6 | 421 | 4 | AAE85042 | AAe85042 Human SER |
| 25 | 2287 | 68.6 | 421 | 8 | ADH17442 | Adh17442 Human NOV |
| 26 | 2254 | 67.6 | 419 | 7 | ADE29367 | ADe29367 Human ser |
| 27 | 2219.5 | 66.6 | 486 | 5 | AAE19171 | AAe19171 Human pro |
| 28 | 2207 | 66.2 | 491 | 4 | AAE93442 | AAe93442 Human pro |
| 29 | 2130.5 | 63.9 | 412 | 8 | ADH17454 | Adh17454 Human NOV |
| 30 | 2108 | 63.2 | 446 | 6 | ABP56843 | ABp56843 Human den |
| 31 | 2097 | 62.9 | 382 | 8 | ADH17444 | Adh17444 Human NOV |
| 32 | 2059 | 61.8 | 406 | 8 | ADH17436 | Adh17436 Human NOV |
| 33 | 2059 | 61.8 | 406 | 8 | ADH17448 | Adh17448 Human NOV |
| 34 | 2027.5 | 60.8 | 449 | 7 | ADE78981 | ADe78981 Human pro |
| 35 | 2026 | 60.8 | 401 | 8 | ADH17422 | Adh17422 Human NOV |
| 36 | 2026 | 60.8 | 418 | 7 | ADE79008 | ADe79008 Human pro |
| 37 | 2016 | 60.5 | 401 | 8 | ADH17426 | Adh17426 Human NOV |
| 38 | 2012 | 60.4 | 401 | 8 | ADH17428 | Adh17428 Human NOV |
| 39 | 2008 | 60.2 | 451 | 7 | ADE78982 | ADe78982 Human pro |
| 40 | 1906.5 | 57.2 | 375 | 7 | ADE78993 | ADe78993 Human pro |
| 41 | 1842.5 | 55.3 | 370 | 8 | ADH17424 | Adh17424 Human NOV |
| 42 | 1839.5 | 55.2 | 367 | 8 | ADH17420 | Adh17420 Human NOV |
| 43 | 1474 | 44.2 | 283 | 8 | ADH17458 | Adh17458 Human NOV |
| 44 | 1311 | 39.3 | 296 | 3 | AAV72108 | AAv72108 Human ser |
| 45 | 1311 | 39.3 | 372 | 3 | AAV72092 | AAv72092 Human ser |

ALIGNMENTS

| | | |
|----------|-------------------------------|---|
| RESULT 1 | ADH17440 | ADH17440 standard; protein; 569 AA. |
| ID | ADH17440 | |
| AC | ADH17440; | |
| DT | 11-MAR-2004 (first entry) | |
| XX | DE | Human NOV12n protein - SEQ ID 130. |
| XX | NOV; | antidiabetic; anorectic; cardiant; hypotensive; |
| XX | KW | antiatertiosclerotic; anorectic; virucide; antibacterial; fungicide; |
| XX | KW | protozoacide; nootropic; neuroprotective; antiparkinsonian; |
| XX | KW | anticonvulsant; osteopathic; antiarthritic; antiinflammatory; |
| XX | KW | dermatological; antiaspheric; antilipemic; metabolic; diabetes; |
| XX | KW | obesity; infection; anorexia; cancer; cardiovascular; hypertension; |
| XX | KW | atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's; |
| XX | KW | epilepsy; immune; osteoarthritis; haemopoietic; |
| XX | KW | inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis; |
| XX | KW | cell differentiation; proliferation; haemopoiesis; wound healing; |
| XX | KW | angiogenesis; gene therapy; chromosome mapping; tissue typing; |
| XX | KW | pharmacogenomic; human. |
| OS | Homo sapiens. | |
| XX | OS | |
| XX | WO2003093432-A2. | |
| PN | 13-NOV-2003. | |
| XX | PD | |
| XX | 02-MAY-2003; 2003WO-US013690. | |
| PF | 02-MAY-2002; 2002US-0377321P. | |
| XX | PR | |

PR 08-MAY-2002; 2002US-0378730P.
 PR 24-MAY-2002; 2002US-0383075P.
 PR 29-MAY-2002; 2002US-0384044P.
 PR 30-MAY-2002; 2002US-0384215P.
 PR 30-MAY-2002; 2002US-0384295P.
 PR 30-MAY-2002; 2002US-0384297P.
 PR 30-MAY-2002; 2002US-0384327P.
 PR 31-MAY-2002; 2002US-0384352P.
 PR 02-JUL-2002; 2002US-0393333P.
 PR 09-AUG-2002; 2002US-0402154P.
 PR 09-AUG-2002; 2002US-0402171P.
 PR 09-AUG-2002; 2002US-0402204P.
 PR 09-AUG-2002; 2002US-0402205P.
 PR 22-AUG-2002; 2002US-0405175P.
 PR 27-AUG-2002; 2002US-0406129P.
 PR 23-SEP-2002; 2002US-0412954P.
 PR 30-SEP-2002; 2002US-0414975P.
 PR 07-OCT-2002; 2002US-0416611P.
 PR 24-OCT-2002; 2002US-0420851P.
 PR 31-OCT-2002; 2002US-0422547P.
 PR 01-MAY-2003; 2003US-00428275.

XX (CURA-) CURAGEN CORP.

PI Alvarez E, Anderson DW, Boldog FL, Catterton E, Edinger SR,
 PI Fernandes ER, Gerlach VL, Gorman L, Grosse WM, Guo X, Ji W,
 PI Kerdus R, Li L, Macdougall JR, Padigaru M, Paturajan M;
 PI Peterson JD, Rastelli L, Shinkens RA, Spletter KA, Stone DJ;
 PI Vermet CAM, Voss EZ, Zhong M;

XX WPI; 2004-053040/05.

DR N-PSDB; ADH17439.

PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
 PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
 PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

XX Claim 1; SEQ ID NO 130; 478bp; English.

CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates antidiabetic, anorectic,
 CC cardiant, hypotensive, antiarteriosclerotic, anorectic, virucide,
 CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
 CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic,
 CC antiinflammatory, dermatological, antiasthmatic and antilipemic
 CC activities. The polypeptides, nucleic acid molecules and antibodies may
 CC be useful in the manufacture of a medicament for treating metabolic
 CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
 CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
 CC diseases including hypertension and atherosclerosis, neurodegenerative
 CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
 CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
 CC acids and polypeptides may also be used as targets for the identification
 CC of small molecules that modulate or inhibit neurogenesis, cell
 CC differentiation, cell proliferation, haemopoiesis, wound healing and
 CC angiogenesis, in gene therapy and the in generation of antibodies or
 CC bind immunospecifically to NOVX substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids may be further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine and pharmacogenomics. The current sequence is that of the human
 CC NOVX protein of the invention.

XX Sequence 569 AA:

Alignment Scores:

Pred. No.: 4,19e-181 Length: 569
 Score: 3004.00 Matches: 563
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 90.1% Indels: 0
 DB: 8 Gaps: 0

US-10-806-370-11 (1-1748) x ADH17440 (1-569)

| | | | |
|----|------|---|------|
| QY | 8 | ACCATGGAGAGGACACCCACGGGAATGATCTCTCCAGCAAGAAACCTTCAGCTGAGCA | 67 |
| DB | 4 | ThrMetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAla | 23 |
| QY | 68 | TCTCCAGCCCGGATCTCCAGCTGGGACACCTCCAGGCGGAGCATCTCCAGCCAGCA | 127 |
| DB | 24 | SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla | 43 |
| QY | 128 | TCTCCAGCCCGGATCTCCAGCTGGGACACCTCCAGGCGGAGCATCTCCAGCCAGCA | 187 |
| DB | 44 | SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla | 63 |
| QY | 188 | TCTCCAGCTGTATACCTCCAGGCGGAGCATCTCCAGGCGGAGCATCTCCAGCCAGCA | 247 |
| DB | 64 | SerProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAla | 83 |
| QY | 248 | TCTCCAGCCCGGATCTCCAGCTGGGACACCTCCAGGCGGAGCATCTCCAGGAGTCA | 307 |
| DB | 84 | SerProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSer | 103 |
| QY | 308 | TCATCCCGCAGCTGAGCTGGTGACAACTCCCAACAGAGTGATCTTTAGACA | 367 |
| DB | 104 | SerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAla | 123 |
| QY | 368 | ACACGAGTGGGGGGTACCACCATCCGATCTCTCCAGCTGGGACAGCAACAGAG | 427 |
| DB | 124 | ThrProValGlyAlaValProLeuArgSerSerProAlaArgSerAlaProAlaThrArg | 143 |
| QY | 428 | GCCACGAGGAGAGCCAGGATACAGGCTGCCAAGTTCACTGGCGGAGGCGAGAG | 487 |
| DB | 144 | AlaThrArgGluSerProGlyThrThrSerLeuProLysSerThrArgGlyGlnLys | 163 |
| QY | 488 | CAGTTACCGCTCATCGGGTGGCTCTCTCTCAATTCCTGGTGGTGGTCTATCATC | 547 |
| DB | 164 | GlnLeuProLeuLeuGlyCysValLeuLeuLeuLeuAlaLeuValAlaSerLeuLeu | 183 |
| QY | 548 | CTCTTCAGTTCTGGGAGGCGCACAGAGGATCAGTCAAGAGGAGGAGGAGAGCTGT | 607 |
| DB | 184 | LeuPheGlnPheTrpGlnGlyHisThrGlyLeuArgTyrGluGlnAlaGluSerCys | 203 |
| QY | 608 | CCCAAGACAGCTGTCTCCCTGTGACGGGGTGGTGGTGGTCAAGCTGAAGTGAAGCTG | 667 |
| DB | 204 | ProLysHisAlaValArgCysAspGlyValValAspCysLysLeuLysSerAspGlnLeu | 223 |
| QY | 668 | GGCTGGCTGAGTTTGACTGGGACAAAGTCTGTAAATCTACTGGTCTCCCAT | 727 |
| DB | 224 | GlyCysValArgPheAspTrpAspLysSerLeuLysLysLysSerGlySerSerHis | 243 |
| QY | 728 | CAGTGGCTCCCATCTGTACAGCAACTGGAAATGACTCTTACAGAGAAAGCTGCGAG | 787 |
| DB | 244 | GlnTrpLeuProLysCysSerSerAsnTrpAsnAspSerGlyLysThrCysGln | 263 |
| QY | 788 | CAGCTGGTTTTCAGAGTGTCAACCGGACAAACGAGTTGGCCACAGGATTTTCCAAAC | 847 |
| DB | 264 | GlnLeuLysGlnGluSerAlaHisArgThrThrGluValAlaHisAlaGlyAspPheAlaAsn | 283 |
| QY | 848 | AGCTTTCGAATCTTGAATACAACTCCACCAATCCAGAAAGCTTCCACAGCTGAAATGC | 907 |
| DB | 284 | SerPheSerLeuLeuArgTyrAsnSerThrLeuGlnLeuSerLeuHisArgSerGlnCys | 303 |
| QY | 908 | CGTTCCAGCGGATATATCTCCCTCAATGTTCACAGTCCGAGTGAAGGAGCATGACCGGG | 967 |
| DB | 304 | ProSerGlnArgTyrLysSerLeuGlnCysSerHisCysGlyLeuAlaMetThrGly | 323 |
| QY | 968 | CGGATCGTGGAGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCAC | 1027 |
| DB | 324 | ArgLeuValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHis | 343 |
| QY | 1028 | TTGGGACACCAACCAATCTGTGAGGACGCTCATTTAGCGCCAGTGGGTGCTACTGGCC | 1087 |

| | | | |
|----------|-------------------------------|--|------|
| D | 344 | Phegilythrlthrlthrlstlelscylgyllythrlleuileasrlalaglntprvalleuthrla | 363 |
| Q | 1088 | GGCCACGCTCTTCTTGTAACCCGGGAGAAGCTCTGAGGGCTGGAGGCTGACCGGAC | 1147 |
| D | 364 | AlahlsCysPhepnevalthrlArgslulysValleuglulglyTrpLysValIlyrLaglly | 383 |
| Q | 1148 | ACCGAACACCGACCGACGAGTGGCTGGAGGACGCTTCATGGCCGACATCATCATCAACG | 1207 |
| D | 384 | ThlserleuileuHlsglileuProglulalalaserllealaglullelleleasner | 403 |
| Q | 1208 | AATPAACACCGATGAGAGAGACGACTATGACATCGGCTCATGCGGCTGTCACAGCCCTG | 1267 |
| D | 404 | AsnlythrlthrlpArglulaspAsprrytrAsprrllealeuMetArghuSerlyserleu | 423 |
| Q | 1268 | ACCCTGTCCGTCACATCCACCCCTGCTTCCCTCCCATGACATGACACAGACCTTAACTC | 1327 |
| D | 424 | ThrlseuSerlAlahlsllleHlelbrlAlaCysleuPromethlsglylglntHrpheserleu | 443 |
| Q | 1328 | AATGAGACCTGCTGGATCACAGGCTTTGGCAAGACACGAGAGACAGATGACAAACATCC | 1387 |
| D | 444 | AsnglulthrlCysrrlrllethrlgllyPhegllylstrArglulthrlAspAspLythrSer | 463 |
| Q | 1388 | CCCTTCGCCGGGAGGTCCAGGTCAATCGATTCGACTTCAGAAATGCAATGACTTACTTG | 1447 |
| D | 464 | ProPheleuArglulValglulAlaAsnleuileasPheLysCysHsnAsprryrlu | 483 |
| Q | 1448 | GTCATATGACATTAACCTTACCCCAAGATGATGTGCTGGGACCTTGCTGGGGCAGA | 1507 |
| D | 484 | VallyrthAspserlyrleuThrlProArghmetetCysAlaglYAspLeuArgllylArg | 503 |
| Q | 1508 | GACTCTGCCAGGAGACAGCGGGGGCTCTTGTCTGTAGACAGACACCGCTGGTAC | 1567 |
| D | 504 | AspserCysglnglYAspserClYglYProleuValCysglnglAlaAsnArgrryrr | 523 |
| Q | 1568 | CTGGCAGGTGCACACAGCTGGGGCACAGCTGTGGCCAGAGAAACAACCTGGTGTAC | 1627 |
| D | 524 | LeuAlaglYAlthrlserTrpGlYthrlglCysglYglAlaArgAsnLysProglYAllyr | 543 |
| Q | 1628 | ACCAAGTGACAGAACTTCTCCCTGATTTTACAGCAAGATGAGAGCGGTGCATTTC | 1687 |
| D | 544 | ThrlsValIthrglulValleuProtrlpleYrlyserlysmetGluserClulValArgPhe | 563 |
| Q | 1688 | AGAAATCC 1696 | |
| D | 564 | Arglysser 566 | |
| RESULT 2 | | | |
| AAE17238 | | | |
| ID | AAE17238 | standard; protein; 562 AA. | |
| XX | AAE17238; | | |
| AC | | | |
| XX | | | |
| XX | 18-APR-2002 | (first entry) | |
| DT | | | |
| XX | | | |
| DE | | Human transmembrane serine protease. | |
| XX | | | |
| KM | | Human: transmembrane serine protease; gene therapy; metastasis; tumour; | |
| KM | | chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation; | |
| KM | | atherosclerosis; neurodegenerative disease; neuroprotective; cytostatic; | |
| KM | | pathogenic infection; antiinflammatory; antiarteriosclerotic; | |
| KM | | antibacterial. | |
| XX | | | |
| OS | | Homo sapiens. | |
| XX | | | |
| XX | WO200196538-A2. | | |
| FN | | | |
| XX | | | |
| PD | 20-DEC-2001. | | |
| XX | | | |
| PF | 12-JUN-2001; 2001WO-EP006618. | | |
| XX | | | |
| PR | 13-JUN-2000; 2000US-0211224P. | | |
| PR | 13-APR-2001; 2001US-0283353P. | | |
| PR | 16-APR-2001; 2001US-0283648P. | | |

| | |
|--|---|
| XX | (FARB) BAYER AG. |
| PA | Xiao Y, Gedrich R; |
| PI | WPI; 2002-098065/13. |
| XX | N-PSDB; AAD7734. |
| DR | |
| XX | Novel isolated polynucleotide encoding transmembrane serine protease |
| PT | polypeptide, for treating chronic obstructive pulmonary disease, tumor |
| PT | angiogenesis, inflammation, atherosclerosis and neurodegenerative |
| XX | disease. |
| XX | |
| XX | Claim 1; Fig 1; 120pp; English. |
| CC | The present invention relates to an isolated polynucleotide encoding a |
| CC | transmembrane serine protease polypeptide. Transmembrane serine protease |
| CC | gene is useful in gene therapy. The invention also relates to a |
| CC | pharmaceutical composition which is useful for modulating the activity of |
| CC | transmembrane serine protease in a disease, such as chronic obstructive |
| CC | pulmonary disease (COPD), metastasis of malignant cells, tumour |
| CC | angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease |
| CC | or pathogenic infection. Transmembrane serine protease is useful as a |
| CC | bait protein in a two-hybrid or three-hybrid assay. The polypeptide is |
| CC | useful for generating antibodies against it and in various assay systems |
| CC | The present sequence is a human transmembrane serine protease |
| XX | |
| SQ | Sequence 562 AA: |
| Alignment Scores: | |
| Pred. No.: | 8.66e-181 Length: 562 |
| Score: | 2999.00 Matches: 562 |
| Percent Similarity: | 100.0% Conservative: 0 |
| Best Local Similarity: | 100.0% Mismatches: 0 |
| Query Match: | 90.0% Indels: 0 |
| DB: | Gaps: 0 |
| US-10-806-370-11 (1-1748) x AAE17238 (1-562) | |
| OY | 11 ATGAGGAGGACAGCCAGCGAATGCATCTCCAGCAAGAACCCTTCAGTGGAGCATCT 70 |
| Db | 1 MerglunrghapserhisielgylashmlaserProhlaahlygmProserAlglYAlaser 20 |
| OY | 71 CAGGCCAGGCATCTTCAGCTGGACAACCTCCAGGCCGGGACATCTCCAGCCAGGCATCT 130 |
| Db | 21 ProhlaglnhlaserProhlaglythrProprogilYarGalaserProhlaglnhlaser 40 |
| OY | 131 CAGGCCAGGCATCTTCAGCTGGAGACACTCCGGGCCGGGACATCTTCAGGCCAGGCATCT 190 |
| Db | 41 ProhlaglnhlaserProhlaglyThrProprogilYarGalaserProhlaglnhlaser 60 |
| OY | 191 CAGCTGTACACCTCCAGCGGGGACATCCAGGCCGGGACATCCAGCCAGGCATCT 250 |
| Db | 61 ProhlaglyThrProprogilYarGalaserProgilYarGalaserProhlaglnhlaser 80 |
| OY | 251 CAGCCCGGGCATCTCCGGCTCTGGGCATCACTTCAGGTCCTCATCCGGAGGTCATCA 310 |
| Db | 81 ProhlahrghlaserProhlaehmlaserleuserrahrgseSerSerClYArgrSerSer 100 |
| OY | 311 TCCGCCAGGTACAGCTCGGTGAACAACCTCCCAACCAAGTGATACCTTTGTAAAGCAACA 370 |
| Db | 101 SerAlargSerlalaserValThmrhrSerProThmrArgvalTYLeuValArghlAthr 120 |
| OY | 371 CCAGTGGGGCTGTACCATCCGATCATCTCTCGGAGGTACAGCACGCAACACAGGGCC 430 |
| Db | 121 ProvalglYAlaAlaProhleargserSerProhlaahgSerAlaProhlarThrhghla 140 |
| OY | 431 ACCAGGAGAGCCAGGTAAGACTGCCCCAAGTTCACCTGGCGGAGGACCCAGAACGAG 490 |
| Db | 141 ThrArghluserProclYthrSerleuProuysPhetrnTrpArghlGlYglhlgln 160 |
| OY | 491 CTACCGCTATCGGGTGGCTGCTCTCTCATATTGCCCTGGTGGTTTGCTCATCATCTC 550 |

Db 161 LeuProLeu11eG1yCyVa11euleuLeu11eal1eulVal1SerLeu11e1leu 180
 Qy 551 TTCCAGTTCTGGCAGGCGCACAGAGATCAGTACAGAGAGAGAGAGAGTGTCCC 610
 Db 181 PheGlnPheTrpGlnGlyHisThrGly11eArGTrpLysGlnGlnArgGlnSerCyber 200
 Qy 611 AAGCAGCGTGTGGCTGTGACGGGGGTGTGACATGCAAGCTGAAAGAGTGAAGTGGCC 670
 Db 201 LysHisAlaValArgCyAspGlyValAlaAspCyLysLeuLysSerAspGlnLeuGly 220
 Qy 671 TGGGAGAGGTGACTGGGAGCAAGCTGTCTTAAATCTACTCTGGGCTCTCCCATAG 730
 Db 221 CysValAlaGlnPheAspTrpAspLysSerLeuLeuLysIleTrpSerGlySerSerHisGln 240
 Qy 731 TGGCTTCCATCTGTAGCAGCACTGGAATGACTCTTACTAGAGAAAGCTGGCAGAG 790
 Db 241 TrpLeuProIleCySerSerSerHisTrpAsnAspSerTrpSerGlyLysTrpCyGlnGln 260
 Qy 791 CTGGGTTTCCAGAGTCTCACCCGACAAACGAGGTTGCCACAGGATTTTGCACAGC 850
 Db 261 LeuGlyPheGlnSerAlaHisArgThrTrpGlyValAlaHisArgAspPheAlaAsnSer 280
 Qy 851 TTTCATCTTGAATACAACTCCACCAATCCAGAAAGCTCCACAGGCTCGAATGCCCT 910
 Db 281 PheSerIleLeuArgTrpAsnSerThrIleGlnGlnSerLeuHisArgSerGlnCyber 300
 Qy 911 TCCGAGCGATATCTCCCTCAGTGTTCCTCAGTCCGAGTCCAGGAGCCATGACCCGCG 970
 Db 301 SerGlnArgTrpIleSerLeuGlnCySerHisCySerGlyLeuArgAlaMetThrGlyArg 320
 Qy 971 ATCGTGGAGGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAATCTGCATT 1030
 Db 321 IleValGlyGlyAlaLeuAlaSerAspSerTrpProTrpGlnValSerLeuHisPhe 340
 Qy 1031 GGCACACCCCATCTGTGAGAGCAGCTCATTAAGCCGCCAGTGGGTCTCACTGCCGCC 1090
 Db 341 GlyThrThrHisIleCySerGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
 Qy 1091 CACTGCTTCTTCTGACCCCGGAGAGAGTCTTGGAGGGCTGGAAGGTGTACCGCGGAC 1150
 Db 361 HisCySerPheValIleTrpArgGlnLysValLeuGlnGlyTrpLysValTyAlaGlyThr 380
 Qy 1151 AGCAACTGCAACCGTTGCTGAGGAGAGCTTCCATGGCCGATCATCATCAACGCAAT 1210
 Db 381 SerAsnLeuHisGlnLeuProGlnAlaLeuSerIleAlaGlnIleIleAsnSerAsn 400
 Qy 1211 TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGGGCTGCCAAGCCCTGAC 1270
 Db 401 TyrThrAspGlnGlnAspAspTrpAspIleAlaLeuMetArgLysLeuSerProLeuThr 420
 Qy 1271 CTGTCCGCTCATCCACCCCTGCTTGCCTCCCATGATGACAGACCTTATGCTCAAT 1330
 Db 421 LeuSerAlaHisIleHisProAlaCyLeuProMetHisGlnGlnIleThrPheSerLeuAsn 440
 Qy 1331 GAGACCTGCTGGATCAACGGCTTGGCAACACAGGAGACAGATGACAGACATCCCC 1390
 Db 441 GlnThrCySerTrpIleTrpGlyPheGlyLysThrArgGlnTrpAspAspLysThrSerPro 460
 Qy 1391 TTCTCCGGGAGGAGGAGGATCTCATGACTTCAAGAAATGCAATGACTACTGTGTC 1450
 Db 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysCyAsnAspTrpLeuAla 480
 Qy 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGCTTCTGTGGGAGAGAC 1510
 Db 481 TyrAspSerTrpLysLeuThrProArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
 Qy 1511 TCTGCGCAGGAGACAGCGGGGGGCTTTGTCTGTGAGCAAGAACCGTGGTACTGT 1570
 Db 501 SerCySerGlnGlyAspSerGlyGlyProLeuValCySerGlnGlnAsnArgTrpTrpLeu 520
 Qy 1571 GCAGGTGACACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAAACCGGTGTATCACCC 1630
 Db 521 AlaGlyValIleThrSerTrpGlyThrGlyCySerGlyGlnArgAsnLysProGlyValIleThr 540

Qy 1631 AAAGTGAAGAGTTCTTCCCTGGATTATACGCAAGATGAGAGCGAGTGGATTGACA 1690
 Db 541 LysValThrGlnValLeuProTrpIleTrpSerLysMetGlnSerGlnValArgPheArg 560
 Qy 1691 AAATCC 1696
 Db 561 LysSer 562
 RESULT 3
 ID AA01943 standard; protein: 562 AA.
 AC AA01943;
 DT 31-JUL-2001 (first entry)
 DE Human transmembrane serine protease (Endothelinase 2-8) protein.
 KW Human; endothelinase 2-8; protease domain; cytosolic; vulnary; wound;
 KW noctropic; peridontitis; dermatological disorder; gene therapy; scar;
 KW angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
 KW chronic inflammatory disease; ocular disorder; circulatory disorder;
 KW crest syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
 KW liver cirrhosis; osteoradionecrosis; systemic sclerosis; osteopagial;
 KW inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
 KW systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
 KW transmembrane serine protease.
 OS Homo sapiens.
 FX Key Location/Qualifiers
 FT Domain 321..562
 FT /label= Protease_domain
 PN MO200136604-A2.
 XX 25-MAY-2001.
 PD 17-NOV-2000; 2000MO-UE031803.
 PF 18-NOV-1999; 9GUS-0166391P.
 PR 22-SEP-2000; 2000US-0234840P.
 XX (CORV-) CORVAS INT INC.
 PA Madison BL, Ong EO;
 PI WPI; 2001-336001/35.
 DR N-PSDB; AAD05796.
 XX New nucleic acid encoding a protein comprising endothelinase activity,
 PT useful in the prevention and treatment of e.g. vascular malformations,
 PT cardiovascular disorders, and chronic inflammatory disease.
 PS Claim 4; Page 138-139; 152pp; English.
 The present sequence is human short form transmembrane serine protease
 (Endothelinase 2-8) protein. Endothelinase 2 is a type-II membrane-type
 serine protease which has a transmembrane domain at the N-terminus,
 followed by a single low density lipoprotein-A receptor domain and a
 single scavenger-receptor cysteine-rich domain. The C-terminus of the
 endothelinase 2 contains the trypsin-like serine protease domain
 characterised by the catalytic triad residues in 3 highly conserved
 regions of the catalytic domain. In addition 3 repetitive sequence are
 found just before the transmembrane domain and represent a sequence motif
 for N-myristoylation modification. The invention relates to an
 endothelinase protein, endothelinase protease domain and their
 corresponding nucleic acid molecules. An endothelinase protein or protease
 domain of it is useful for the treatment and diagnosis of disorders
 associated with aberrant angiogenesis or undesired neovascularisation.
 The undesired angiogenesis is associated with disorders selected from
 solid neoplasm, vascular malformations and cardiovascular disorders such

CC as angiodiroma, angiolipoma, atherosclerosis, restenosis/reperfusion
CC injury, arteriovenous malformations, haemangiomatosis and vascular
CC adenoma, dyschondroplasia with vascular hamartomas (Pafucci's
CC syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
CC syndrome) and Von Hippel Lindau syndrome, chronic inflammatory diseases
CC such as diabetes mellitus, haemophilic joints, inflammatory bowel
CC disease, nonhealing fractures, periodontitis, psoriasis, rheumatoid
CC arthritis, venous stasis ulcers, granuloma-in-burns, hypertrophic scars,
CC liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
CC granuloma and systemic sclerosis and aberrant wound repair, circulatory
CC disorders Raynaud's phenomenon, cret syndromes such as calcinosis,
CC oesophagial, dyomecillopy, sclerodactyly and teangiectasis, dermatological
CC disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
CC vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine
CC stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome
CC and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
CC caused by ocular neovascular disease, corneal graft neovascularization,
CC macular degeneration, retinopathy of prematurity, retrolental fibroplasia
CC and corneal neovascularisation. The nucleic acids of the invention are
CC also used in gene therapy. The invention also provides method for
CC screening compounds that modulate angiogenesis

XX Sequence 562 AA:

Alignment Scores:
Pred. No.: 2.77e-180 Length: 562
Score: 2991.00 Matches: 561
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 89.7% Indels: 0
DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x AAE01943 (1-562)

QY 11 ATGGAGAGGAGCAGACGCGGATGATCTTCACGAAAGAACCTTCAGTGCATCT 70
Db 1 MetelunrghasberhlsghlyashalaserProhlaarghtrProserAlglYAlaser 20
QY 71 CCAGCCGAGGATCTCCAGCTGGGACACCTTCAGGCGGCGCATCTCCAGCCGAGCATCT 130
Db 21 ProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 40
QY 131 CCAGCCGAGGATCTCCAGCTGGGACACCTTCAGGCGGCGCATCTCCAGCCGAGCATCT 190
Db 41 ProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 60
QY 191 CCAGCTGGTACACCTTCAGGCGGCGCATCTTCAGGCGGCGCATCTTCAGCCGAGCATCT 250
Db 61 ProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 80
QY 251 CCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGCGCATCTTCAGCCGAGCATCT 310
Db 81 ProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 100
QY 311 TCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGCGCATCTTCAGCCGAGCATCT 370
Db 101 SerilaarghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 120
QY 371 CCAGTGGGCGGTGACCATCTTCAGCTTCCTTCAGTGCATCTTCAGCCGAGCATCT 430
Db 121 ProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 140
QY 431 ACAGGAGAGAGCCAGTACAGAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 490
Db 141 ThraghtrProhlaarghtrProhlaarghtrProhlaarghtrProhlaarghtr 160
QY 491 CTACCGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 550
Db 161 LeuproleuilegYCyvalleuileuileuileuileuileuileuileuileuileuileu 180
QY 551 TTCAGTTCAGGAGGAGCAGACGAGTACAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAG 610
Db 181 PheglnpethrpglnghlyHlshtllytlearglylsglnghlnghlnghlnghlnghlngh 200

QY 611 AACGAGCTGTTCGTCGTGACGCGGAGTGTGACCTGCAGCTGAAGATGACGAGCTGGC 670
Db 201 LysHlsHlsAlaValrGysAspGlyValAlaAspCyAlaYlsleuYsSerAspGlnleuGly 220
QY 671 TGGCGAGGTTTGTAGTGGAGCAAGCTTCCTTAAATCTACTCTGGGCTTCGCATCAG 730
Db 221 CyValAlarGhAspTrpAspYlsSerleuileuileuileuileuileuileuileuileuileu 240
QY 731 TGGCTTCCTTCGTTCGACAGCAACTGCATGATCTTCATTCAGAGAGAGAGAGAGAGAG 790
Db 241 TrpLeuProleuYsSerAspTrpAspYlsSerleuileuileuileuileuileuileuileu 260
QY 791 CTGGGTTTGGAGGTGTACCGGAGCAAGAGGAGGTTTCCACAGGAGTTTTCGCAAGCAG 850
Db 261 LeuGlyPheleuSerAlaHlsAlaGlnTrhGlyValAlaHlsAspAspPheHlsAspSer 280
QY 851 TTCATCTTGTAGATCAACCTCCAGCATCCAGAAAGCTTCACAGGTGTGAATGCTCTC 910
Db 281 PheSerleuileuArgTrpAspSerTrhIleGlnGlnSerleuHlsAspSerGlnCyPro 300
QY 911 TCCAGCGGTATATCTTCCTTCAGGTTCCTCCAGCTGCGGATGAGGCGCATGACCGGCGG 970
Db 301 SerGlnArgTrpIleSerleuGlnCySerHlsCySgIyLeuArgAlaMetTrhGlyArg 320
QY 971 ATGCTGGAGGAGGCGGCTGCGGATGAGCAAGTGGCTTCGCAAGTGAAGTCTGACCTTC 1030
Db 321 IleValIlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 340
QY 1031 GGCACCAACCAATCTGTGAGAGGACGCTCATTTGACCCGAGTGGGTGCTCAGTCCGCC 1090
Db 341 GlyTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrhTrh 360
QY 1091 CACTGCTTCTTCTGTCACCCGAGAGAGGTCCTGAGGCTGGAAAGTGTACCGGCGGAC 1150
Db 361 HlsCySphPheValAlaTrhArgGlyValleuIleuIlyTrpYsValYAlaGlyThr 380
QY 1151 AGCACTTCGACAGCTTCGAGGAGGCTTCATTCGAGGATCATCATCAAGCAT 1210
Db 381 SerAsnleuHlsGlnleuProGlnAlaAlaSerleuIleAlaGlnleuIleleuSerHls 400
QY 1211 TACACGATGAG 1270
Db 401 TyrTrhAspGlnGlnAspAspTrpAspIleAlaAlaMetArgleuSerYsProleuThr 420
QY 1271 CTGTCCCTCAGATCCACCTGCTTCCTTCCTCCATGATGACAGAGACTTTCAGCTCAAT 1330
Db 421 LeuSerAlaHlsIleHlsProAlaCyAlaProMetHlsGlyGlnTrhPheSerleuAsn 440
QY 1331 GAGACCTGCGGATCAAGAGGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390
Db 441 GluTrhYsTrpIleTrhGlyPheGlyYsTrhArgGlnTrhAspAspYsTrhSerPro 460
QY 1391 TTCCTCCGAGAGGAG 1450
Db 461 PheleuArgGlnValGlnValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 480
QY 1451 TATGACATTAATCTTACCCCAAGATGATGTGTCTGGGAGCTTCGTTGGGCGGAGAGAC 1510
Db 481 TyrAspSerTrpLeuTrhProArgMetCysAlaGlyAspLeuValArgGlyAlaArgAsp 500
QY 1511 TCCGCGAGGAG 1570
Db 501 SerCySglnGlyAspSerGlyGlyProleuValCySglnGlnAsnAsnAspTrpYsleu 520
QY 1571 GAGGTGTACACAGCTGGGAG 1630
Db 521 AlaGlyValTrhSerTrpGlyTrhGlyCySgIyGlnTrhAsnYsProGlyValYTrhTrh 540
QY 1631 AAAGTGAAGAGTTTCTTCTTCGATTTTACAGCAAGATGAGAGAGAGAGAGAGAGAGAGAG 1690
Db 541 LysValTrhGlnValleuProTrpIleTrpYsSerYsMetGlnSerGlnValAlaArgPheIle 560

QY 1691 AAATCC 1696
|||||
Db 561 LysSer 562

RESULT 4

AD110393
ID AD110393 standard; protein; 562 AA.

XX AD110393;

DT 22-APR-2004 (first entry)

XX Human cell surface protease #12.

XX therapeutic agent; plasmin; protease specific antigen; PSA;

XX cell-surface protease-associated disease; cancer; ocular disease;

XX cardiovascular disease; chronic inflammatory disease; wound;

XX circulatory disorder; dermatological disorder; rheumatoid arthritis;

XX psoriasis; diabetic retinopathy; pterygium;

XX excimer laser surgery scarring; glaucoma filtering surgery scarring;

XX macular degeneration; crest syndrome; solid neoplasm; vascular tumour;

XX melanoma; Kaposi's sarcoma; human; cell surface protease.

XX Homo sapiens.

XX MO20295007-A2.

XX 28-NOV-2002.

XX 23-MAY-2002; 2002MO-US016819.

XX 23-MAY-2001; 2001US-0293267P.

XX (CORV-) CORVAS INT INC.

XX Madison BL, Semple JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV;

XX WPI; 2003-221280/21.

XX DR N-PSDB; AD110392.

XX PT Novel conjugate useful for treating cell-surface protease-associated

XX disease, comprises a therapeutic agent and a peptidic or nucleic acid

XX substrate linked to it optionally by a peptidic linker.

XX PS Claim 9; SEQ ID NO 24; 581bp; English.

XX CC The invention comprises a conjugate that consists of a therapeutic agent

XX and a peptide substrate (optionally linked via linker). The peptide

XX substrate is proteolytically cleaved by a cell surface protease pr a

XX soluble, released or shed form of it, to liberate the therapeutic agent,

XX the conjugate of the invention is not substantially cleaved by plasmin or

XX CC protease specific antigen (PSA). The conjugate of the invention is useful

XX for treating a cell-surface protease-associated disease such as: cancer;

XX ocular diseases, cardiovascular diseases, chronic inflammatory diseases,

XX wounds, circulatory disorders, dermatological disorders, rheumatoid

XX arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,

XX CC scarring from excimer laser surgery, scarring from glaucoma filtering

XX CC surgery, macular degeneration, crest syndromes, solid neoplasms, vascular

XX CC tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence

XX represents a human cell surface protease.

XX SQ Sequence 562 AA;

XX Alignment Scores:

XX Pred. No.: 2.77e-180 Length: 562

XX Score: 2991.00 Matches: 561

XX Percent Similarity: 99.8% Conservative: 0

XX Best Local Similarity: 99.8% Mismatches: 1

XX Query Match: 89.7% Indels: 0

XX DB: 7 Gaps: 0

XX US-10-806-370-11 (1-1748) X AD110393 (1-562)

QY 11 ATGAGAGAGGACAGCCAGGGAATGATCTTCAGCAAGAAACCTTCAGTGCATCT 70
Db 1 MetGluArgAspSerHisAGLYAsenHisSerProHisArgThrProSerHisAGLYAsen 20
QY 71 CCAGCCCAAGCATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 130
Db 21 ProHisArgHisSerProHisArgThrProProHisArgHisSerProHisArgHisSer 40
QY 131 CCAGCCCAAGCATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 190
Db 41 ProHisArgHisSerProHisArgThrProProHisArgHisSerProHisArgHisSer 60
QY 191 CCAGCTGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCCAAGCATCT 250
Db 61 ProHisArgThrProProHisArgHisSerProHisArgHisSerProHisArgHisSer 80
QY 251 CCAGCCCGGGGATCTCCAGGCTGGGATCTCCAGCTTCAGGCTTCATCCAGGAGTCTATCA 310
Db 81 ProHisArgHisSerProHisArgHisSerHisSerHisSerHisSerHisSerHisSer 100
QY 311 TCCGCGAGGTCAGGCTCGGTGACCAACCTCCCAAGAGTACCTGTGTAGAGCAACA 370
Db 101 SerHisArgSerHisSerHisSerHisSerHisSerHisSerHisSerHisSerHisSer 120
QY 371 CCAGTGGGGGCTGTACCCATCCATCCATCTCTGCGAGGTACAGCAAGCAAGCAAGGCCC 430
Db 121 ProValHisValHisValHisValHisValHisValHisValHisValHisValHisVal 140
QY 431 ACCAGGAGAGCCAGGATACAGGCTCCCAAGTTCACCTGGCGGGAGGCGGAGCAAGCAG 490
Db 141 ThrArgHisSerProHisArgHisSerHisSerHisSerHisSerHisSerHisSerHis 160
QY 491 CTACCGCTCATCGGAGTCCGCT 550
Db 161 LeuProLeuHisLeuGlyCysValHisLeuLeuHisValHisValHisValHisValHis 180
QY 551 TTCAGGTTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 610
Db 181 PheGlnPheThrPheGlnHisHisHisHisHisHisHisHisHisHisHisHisHisHis 200
QY 611 AAGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 670
Db 201 LysHisValHisValHisValHisValHisValHisValHisValHisValHisValHis 220
QY 671 TGCCTGAGGTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 730
Db 221 CysValHisArgHisArgHisArgHisArgHisArgHisArgHisArgHisArgHisArg 240
QY 731 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790
Db 241 ThrLeuProHisLeuHisSerHisSerHisSerHisSerHisSerHisSerHisSerHis 260
QY 791 CTGGGTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
Db 261 LeuHisValHisValHisValHisValHisValHisValHisValHisValHisValHis 280
QY 851 TTTCATCTTGAAGTACCACTCCAGCATCCAGGAAAGCTCCAGAGTCTGAATGCGCT 910
Db 281 PheSerHisLeuHisArgHisArgHisArgHisArgHisArgHisArgHisArgHisArg 300
QY 911 TCCAGGCGGTATATCTCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
Db 301 SerGlnArgThrHisSerHisSerHisSerHisSerHisSerHisSerHisSerHisSer 320
QY 971 ATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1030
Db 321 HisValHisValHisValHisValHisValHisValHisValHisValHisValHisVal 340
QY 1031 GGCACACCAACATCTGTGAGGAGCAGCTGATGACCCAGAGTGGAGTCTCACTGCGGCC 1090
Db 341 GlnThrHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis 360
QY 1091 CACTGCTTCTTCTGACCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1150


```

Db      361 HiscyshpshpValThrArgGluValLeuIuGlyTrpLysValTyrAlaGlyThr 380
QY      1151 AGCACTGACACCGCTTGGCTGAGGAGCGCTCCATTCGCCAGATCATCATCAGACAAAT 1210
Db      381 SerAsnLeuHisGlnLeuProGlnAlaSerIleAlaGluIleIleLeuSerAsn 400
QY      1211 TACACCGATGAGGAGGAGCACTATGACATGCGCTCATGGGGCTGTGCCAGCCCTGACC 1270
Db      401 TyrThrAspGluGlnAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY      1271 CTGTCCGCTCATCATCACCCTGCTTGGCTCCCGCATGAGGAGACACTTATGACCTCAAT 1310
Db      421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY      1331 GAGACCTGCTGATCAGACGCTTTGGCAAGACGAGGAGACAGATGACAGACATCCCTCC 1390
Db      441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysThrSerPro 460
QY      1391 TTCTCTCCGGGAGGTGCGAGGTATCATCTCATGCACTTCAAGAAATGCATGACTATTGTC 1450
Db      461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
QY      1451 TATACAGATTACCTTACCCCAAGATGATGTGTCTGGGAGCTTCGTGGGGGAGAGAC 1510
Db      481 TyrAspTyrLeuThrProArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
QY      1511 TCCTGCGAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGAGAAACAACCGGTGACCTG 1570
Db      501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY      1571 GCAGGTGTCACACCTGAGGAGGAGCGGTGTGTGCGCCAGAGAAACAACCTGTGTGTACAC 1630
Db      521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY      1631 AAAGTGACAGAAAGTCTTCCCTGATTTACAGCAAGATGAGAGCGAGGTGCAATTGAGA 1690
Db      541 LysValThrGluValLeuProTrpIleTyrSerLysMetGlnSerGluValArgPheIle 560
QY      1691 AATATC 1696
Db      561 LysSer 562

```

```

RESULT 5
ADJ46917
ID      ADJ46917 standard; protein; 562 AA.
XX
AC      ADJ46917;
XX
DT      06-MAY-2004 (first entry)
XX
DE      Human transmembrane serine protease (MSP)-related polypeptide #2.
XX
KW      Human; transmembrane serine protease; MSP; cell surface protease;
KW      plasmin; prostate specific antigen; PSA; proliferative disease;
KW      cell-surface protease-associated disease; autoimmune disease;
KW      inflammatory disease; infectious disease; endocrine disease; cancer;
KW      ocular disorder; cardiovascular disorder; chronic inflammatory disease;
KW      wound; circulatory disorder; dermatological disorder; restenosis;
KW      rheumatoid arthritis; psoriasis; diabetic retinopathy;
KW      laser surgery scarring; glaucoma filtering surgery scarring;
KW      macular degeneration; CREST syndrome; bacterial infection; viral disease;
KW      solid neoplasm; vascular tumour; lung; colon; prostate; melanoma;
KW      Kaposi's sarcoma; enzyme.
XX
OS      Homo sapiens.
XX
PN      US2004001801-A1.
XX
PD      01-JAN-2004.
XX
PF      23-MAY-2002; 2002US-00156214.
XX

```

```

PR      23-MAY-2002; 2002US-00156214.
XX
PA      (CORV-) CORVAS INT INC.
XX
PI      Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV;
XX
DR      WPI; 2004-190126/18.
XX
DR      N-PSDB; ADJ46916.
PT      Conjugate useful for treating e.g. cancer, cell-surface protease-
PT      associated diseases, comprising a peptidic substrate or nucleic acid
PT      substrate linked to a therapeutic agent through a linker.
XX
PS      Claim 9; SEQ ID NO 24; 361bp; English.
XX
CC      The invention relates to a conjugate comprising a therapeutic agent and a
CC      peptidic substrate or nucleic acid substrate linked to the agent
CC      optionally through a linker or peptidic linker, where the peptidic
CC      substrate is proteolytically cleaved by a cell surface protease or a
CC      soluble, released or shed form conjugate to liberate the agent and the
CC      conjugate is not substantially cleaved by plasmin or prostate specific
CC      antigen (PSA). The conjugate is useful for treating a disease, which
CC      involves administering a conjugate to a subject, where the disease is
CC      preferably a proliferative diseases or a cell-surface protease-associated
CC      disease. The diseases include autoimmune diseases, inflammatory diseases,
CC      infectious diseases and endocrine diseases. The conjugate is useful for
CC      treating a cell-surface protease-associated disease, which involves
CC      administering a conjugate comprising an agent and a peptidic substrate to
CC      a subject exhibiting symptoms of a cell-surface protease-associated
CC      disorder, where the disease is selected from cancer, ocular disorders,
CC      cardiovascular disorders, chronic inflammatory diseases, wounds,
CC      circulatory disorders, dermatological disorders, restenosis, rheumatoid
CC      arthritis, psoriasis, diabetic retinopathies, scarring from laser
CC      surgery, scarring from glaucoma filtering surgery, macular degeneration,
CC      CREST syndrome, bacterial infections, viral diseases, solid neoplasms and
CC      vascular tumours such as lung cancer, colon cancer, prostate cancer,
CC      melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
CC      serine protease (MSP)-related polypeptide of the invention.
XX
SQ      Sequence 562 AA:

```

```

Alignment Scores:
Pred. No.:      2,77e-180      Length:      562
Score:          2991.00      Matches:      561
Percent Similarity: 99.8%      Conservative: 0
Best Local Similarity: 99.8%      Mismatches: 1
Query Match:      89.7%      Indels:      0
DB:              8      Gaps:      0

```

```

US-10-806-370-11 (1-1748) x ADJ46917 (1-562)
QY      11 ATGAGAGGAGGACGCGGGAATGCATCTCCAGCAAGAACACCTTGAGCATCT 70
Db      1 MeclunArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY      71 CCAGCCGAGCATCTCCAGCTGGGACACCTCCAGCGCGGCGCATCTCCAGCCGAGCATCT 130
Db      21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY      131 CCAGCCGAGCATCTCCAGCTGGGACACCTCCGGGCGGCGCATCTCCAGCCGAGCATCT 190
Db      41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY      191 CCAGCTGTACACCTCCAGGCGCGGCGCATCTCCAGCGCGGCGCATCTCCAGCCGAGCATCT 250
Db      61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY      251 CCAGCCGCGGCGCATCTCCGCTGGGACATCTTCCAGGTCTCATCCGCGAGGTCATCA 310
Db      81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
QY      311 TCCGCGAGGTGAGCTCGGTGAGCAACCTCCCAACGAGGTGATCTGTAGAGCAACA 370

```

Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
 QY 371 CCAGTGGGGGCTGTACCCATCATCTCTGCGAGGTGAGCAGCAACAGGGGCC 430
 Db 121 ProValGlyAlaValProIleArgSerSerProIleArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGAGCCAGTAGAGAGCTGCGCAAGTTACCTGGCGGAGAGCCAGAAAGAG 490
 Db 141 ThrArgIleSerProGlyThrSerLeuProIleThrArgGluGlyGlnGlyGln 160
 QY 491 CTACCGCATCGGGTGGTCTCTCTCATTTGGCCGTGGTGGTGGTGGTGGTGGTGGTGGT 550
 Db 161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValIleSerLeuIleLeu 180
 QY 551 TTCCAGTTCTGGAGGAGGAGCAGAGGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610
 Db 181 PheGlnPheTrpGlnGlyIleThrGlyIleArgGlyIleArgGlyIleArgGlyIleArg 200
 QY 611 AAGACGCTGTGCTGTGACGGAGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 670
 Db 201 LysHlaIaValAlaArgCysAspGlyValIleAlaSerCysLeuLysSerAspGlyLeuGly 220
 QY 671 TGGGTGAGAGTTGAGTGGAGAGAGTCTGCTTAAATCTACTGAGGTCTCTCCATGAG 730
 Db 221 CysValAlaArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHla 240
 QY 731 TGGCTTCCATCTGTGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 790
 Db 241 TrpLeuProIleCysSerSerHlaTrpAsnHlaSerTyrSerGlyLysThrCysGlnGln 260
 QY 791 CTGGGTTTGAAGAGTCTCACCGGAGCAACGAGGTGGCCACAGGATTTTCCAAACAGC 850
 Db 261 LeuGlyPheGlnSerAlaHlaIleArgThrThrGluValAlaHlaIleArgAspPheAla 280
 QY 851 TTCTCAATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 910
 Db 281 PheSerIleLeuAlaGlyTyrAsnSerThrIleGlnGlnSerLeuHlaIleArgSerGly 300
 QY 911 TCCAGCGGTATATCTCCCTCAGTGTCCCATGCGAGTGAAGGAGGAGGAGGAGGAGGAG 970
 Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHlaCysGlyLeuAlaGlnMetThrGlyArg 320
 QY 971 ATCTGGAGAGAGGCGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1030
 Db 321 IleValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHla 340
 QY 1031 GGACACCAACCATCTGTGAGAGAGCGCTCATGACGCGGAGTGGTGGTGGTGGTGGTGG 1090
 Db 341 GlyThrThrHlaIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaIle 360
 QY 1091 CACTGCTTCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1150
 Db 361 HisCysPhePheValThrArgGluLysValLeuGluGlyTrpLysValTyrAlaGlyThr 380
 QY 1151 AGCAACCTGACCAAGTGGCTGAGGAGAGCTTCATGCGAGATCATCAACAGCAAT 1210
 Db 381 SerLeuLeuHlaGlnLeuProGluAlaIleSerIleAlaGluIleIleIleAlaSerAsn 400
 QY 1211 TACACCGATGAGAGAGAGATATGACATGCGCCCTCAACCGGCTGCAACCCCTGAC 1270
 Db 401 TyrThrAspGluGluAspAspLysTrpIleAlaLeuMetArgLysSerLysProLeuThr 420
 QY 1271 CTGTCCGTCATCATCACCCTGCTGCTCCCATGAGTGAAGAGAGAGAGAGAGAGAGAGAG 1330
 Db 421 LeuSerAlaHlaIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 440
 QY 1331 GAGACCTGCTGAGTACAGGCTTTGGCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390
 Db 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysThrSerPro 460
 QY 1391 TTCCTCCGAG 1450
 Db 461 PheLeuArgGluValGlnValAsnLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480

QY 1451 TATGACAGTTACCTTACCCAGAGATGATGTGTGCGGAGCTTCTGTGGGAGAGAGAC 1510
 Db 481 TyrAspSerTyrLeuThrTrpProArgMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
 QY 1511 TCCTGCCAGGAG 1570
 Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGluGlnAsnAsnArgTrpTyrLeu 520
 QY 1571 GCAGGTGACACCACTGGGAG 1630
 Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
 QY 1631 AAAGTGAAGAGAGTCTTCCCTGTGATTTACAGCAAGATGAGAGAGAGAGAGAGAGAG 1690
 Db 541 LysValThrGluValLeuProTrpIleTyrSerLysMetGluSerGluValArgPheIle 560
 QY 1691 AAATCC 1696
 Db 561 LysSer 562
 RESULT 6
 ADY50146 ID ADY50146 standard; protein; 562 AA.
 AC ADY50146;
 DT 19-MAY-2005 (first entry)
 XX
 DE Human endothelialase-2L.
 XX
 KW Antiarthritic; Antirheumatic; Antiosteoporotic; Antidiabetic;
 KW Ophthalmological; Cardiovascular-Gen.; Vulnerary; Antiinflammatory;
 KW Vasoconstrictor; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
 KW rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
 KW dermatological disease; cancer; neoplasm; endothelialase-2; enzyme.
 XX
 OS Homo sapiens.
 XX
 PN MO2005019270-A2.
 PD 03-MAR-2005.
 PF 12-AUG-2004; 2004MO-US026148.
 XX
 PR 14-AUG-2003; 2003US-0495005P.
 PR 14-NOV-2003; 2003US-0520164P.
 XX
 PA (DYAX-) DYAX CORP.
 PA (DEND-) DENDREON CORP.
 XX
 PI Madison EL, Nixon A;
 XX
 DR WPI; 2005-202609/21.
 DR N-PSDB; ADY50145.
 PT Novel protein capable of inhibiting endothelialase-2, useful for treating
 or preventing angiogenesis related disorder e.g. cancer.
 XX
 PS Disclosure; SEQ ID NO 94; 157bp; English.
 XX
 CC The invention relates to an isolated protein (I) which comprises a heavy
 CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
 CC immunoglobulin variable domain sequence, where the first and second
 CC immunoglobulin variable domain sequences from an antigen binding site
 CC that specifically binds to human endothelialase-2. (I) is useful for
 CC detecting an endothelialase or endothelialase activity in a sample, for
 CC modulating an activity of an ET2-expressing cell, for modulating
 CC proteolysis, for killing or inhibiting growth of a cell, for detecting
 CC endothelialase in a subject, for modulating endothelialase activity in a
 CC subject, for treating or preventing a disorder characterized by unwanted
 CC angiogenesis in a subject. The disorder is chosen from rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, ocular disorder such as

KM Angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
 KM chronic inflammatory disease; ocular disorder; circulatory disorder;
 KM crest syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
 KM liver cirrhosis, osteoradionecrosis, systemic sclerosis; oesophageal;
 KM inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
 KM systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
 KM transmembrane serine protease.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 321..688
 FT /label= Protease_domain
 XX MO200136604-A2.
 XX 25-MAY-2001.
 XX 17-NOV-2000; 2000WO-US011803.
 XX 18-NOV-1999; 99US-0166391P.
 XX 22-SEP-2000; 2000US-0234840P.
 XX (CORV-) CORVAS INT INC.
 XX Madison BL, Ong EO;
 XX MPI; 2001-336001/35.
 DR N-PSDB; AAD05797.
 PT New nucleic acid encoding a protein comprising endothelinase activity,
 PT useful in the prevention and treatment of e.g. vascular malformations,
 PT cardiovascular disorders, and chronic inflammatory disease.
 XX Claim 4; Page 142-143; 152pp; English.

The present sequence is human long form transmembrane serine protease (Endothelinase 2-I) protein. Endothelinase 2 is a type-II membrane-type serine protease which has a transmembrane domain at the N-terminus, followed by a single low density lipoprotein-A receptor domain and a single scavenger-receptor cysteine-rich domain. The C-terminus of the endothelinase 2 contains the trypsin-like serine protease domain characterized by the catalytic triad residues in 3 highly conserved regions of the catalytic domain. In addition 3 repetitive sequence are found just before the transmembrane domain and represent a sequence motif for N-myristoylation modification. The invention relates to an endothelinase protein, endothelinase protease domain and their corresponding nucleic acid molecules. An endothelinase protein or protease domain of it is useful for the treatment and diagnosis of disorders associated with aberrant angiogenesis or undesired neovascularisation. The undesired angiogenesis is associated with disorders selected from solid neoplasm, vascular malformations and cardiovascular disorders such as angiofibroma, angiolipoma, atherosclerosis, restenosis/reperfusion injury, arteriovenous malformations, haemangiomas and vascular adhesions, dyschondroplasia with vascular hamartomas (Patucci's syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber syndrome) and Von Hippel Lindau syndrome, chronic inflammatory diseases such as diabetes mellitus, haemophilic joints, inflammatory bowel disease, nonhealing fractures, periodontitis, psoriasis, rheumatoid arthritis, venous stasis ulcers, granuloma-burns, hypertrophic scars, liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic granuloma and systemic sclerosis and aberrant wound repairs, circulatory disorders Raynaud's phenomenon, crest syndromes such as calcinosis, oesophageal, dyomectolysis, sclerodactyly and telegiectasis, dermatological disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum, vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome and Osler-Weber-Rendu syndrome and ocular disorders such as blindness caused by ocular neovascular disease, corneal graft neovascularisation, macular degeneration, retinopathy of prematurity, retrolental fibroplasia and corneal neovascularisation. The nucleic acids of the invention are also used in gene therapy. The invention also provides method for screening compounds that modulate angiogenesis

XX SQ Sequence 688 AA;
 Alignment Scores:
 Pred. No.: 1,67e-178
 Score: 2963.00
 Percent Similarity: 99.6%
 Best Local Similarity: 99.3%
 Query Match: 88.9%
 DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x AAE01944 (1-688)

QY 11 ATGAGAGAGGACAGCCAGGAAATGATCTCCAGCAAGAAACACCTTGAGCTGAGCATCT 70
 Db 1 MetGluArgAspSerHisLeuAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAGCCAGGATCTCCAGCTGGGACACCTTCAGGCCGGGATCTCCAGCCAGGATCT 130
 Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAGCCAGGATCTCCAGCTGGGACACCTTCAGGCCGGGATCTCCAGCCAGGATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAGCTGTACACCTCCAGGCCGGGATCTTCAGGCCGGGATCTCCAGCCAGGATCT 250
 Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAGCCGGGATCTCCGGCTTGGCATCACTTCCAGGTCCTCATCCGGGATCTCA 310
 Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerArgAlaSerSer 100
 QY 311 TCCGCGAGGTCAGCTCGGTGACCAACTCCCAAGGATGTCCTGTTAGAGCAACA 370
 Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValThrLeuValArgAlaThr 120
 QY 371 CCAAGTGGGGGCTGTACCATTCATCTCCTCCAGGTCAGGACCCAGGAGGCC 430
 Db 121 ProValGlyAlaValProLeuArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGCCAGGATGAGGCTTCCAGTTCACCTTGGCGGAGGGCCAGAACAG 490
 Db 141 ThrArgGluSerProGlyThrThrSerLeuProLysPheThrThrArgGluGlnLysGln 160
 QY 491 CTACCGGTCAATCGGGTGGGCTCTCTCTCATTCGCTGGGATGTCATCATCTCT 550
 Db 161 LeuProLeuLeuGlyCysValLeuLeuLeuLeuAlaLeuValAlaSerLeuLeuLeu 180
 QY 551 TTCAGTTCCTGGCAGGAGCCACAGGATCAGGTACAGAGCAGAGAGAGTGTCC 610
 Db 181 PheGlnThrThrProGlnGlnLysAlaThrGlyIleArgGlyLysGlnGlnLysGlnCysPro 200
 QY 611 AAGACGCTGTTCCTGTGTACCGGGGTGTGACCTGCAAGCTGAAAGATGACAGCTGGC 670
 Db 201 LysHisAlaValArgCysAspArgValValAspCysLysLeuLysSerAspArgLysGly 220
 QY 671 TGGGTGAGGTTTGCTGAGCAAGTCTCTGTTAAATTAATTAATCTGGGCTCTCCATCAG 730
 Db 221 CysValArgPheAspThrPhePheLysSerLeuLeuLysAlaLeuLysArgLysSerHisGln 240
 QY 731 TGGCTTCCTCATCTGTACAGCAACTGGAATGACTCTTACTCAGAGAAAGCTGGCAGAG 790
 Db 241 TrpLeuProLeuCysSerSerSerLeuTrpAspSerLysSerGluLysThrCysGlnGln 260
 QY 791 CTGGGTTTCCAGAGAGTCTCACCGGACAAACGAGGTTCCTCCACAGGATTTTCCCAACAG 850
 Db 261 LeuGlyPheGluSerAlaHisAlaArgThrThrGluValAlaHisAlaGlyAspPheAlaAsnSer 280
 QY 851 TTTCATCTTGTAGATCAACTCCACCATCCAGGAAGAGCTCCACAGAGTGTGAATGCCCT 910
 Db 281 PheSerIleLeuArgLysAsnSerThrIleGlnGlnLysSerLeuHisAlaSerGluCysPro 300

```

QY 911 TCCAGCGGATATATCTCCCTCAGTGTTCCTCAGTGCAGGATGAGGGCATGACCGGGCGG 970
DB 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuAlaGlnMetThrGlyArg 320
QY 971 ATCGTGGAGAGGGGCGCTGGCTCGGATAGCAAGTGGCCCTTGGCAAGTGAATCTGCACCTC 1030
DB 321 IleValGlyGlyAlaLeuAlaSerAspSerIleTyrProTArgIleValSerLeuHisPhe 340
QY 1031 GGCACCAACCCATCTGTGGAGGACGCTCATTCAGCCCGCATGGGTGCTCATCTGCCGCC 1090
DB 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTyrValLeuThrAlaAla 360
QY 1091 CACTGCTTCTTCTGACCCCGGAGAGAGGCTCTGAGAGGGCTGGAAGGTCTACCGGGGAC 1150
DB 361 HisCysPhePheValThrArgIleValLeuGlnGlyTyrIleValIleGlyThr 380
QY 1151 AGCAACCTGACACAGTGGCTTGGAGGACGCTTCATTCATTCGCGATCATCATCAGCAGCAT 1210
DB 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleAsnSerAsn 400
QY 1211 TACACCGATGAGAGAGGACTATGACATGCGCCCTCATGCGGCTGTCCAGCCCTGACC 1270
DB 401 TyrThrAspGlnGluAspTyrAspIleAlaLeuMetArgLeuSerIleProLeuThr 420
QY 1271 CTGTCGGCTGACATCCACCCCTTGGCTTGGCTTGGCTGACATGACGACCTTATAGCTCAT 1330
DB 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY 1331 GAGACCTGCTGATACACAGGCTTTGGCAAGACACGAGGAGACAGATGACAGATATCCCC 1390
DB 441 GluThrCysTyrIleThrGlyPheGlyIleThrArgIleThrAspIleThrSerPro 460
QY 1391 TTCTCTCGGAGGCTGACAGTCAATCTCATTCGACTTCAAGAAATGATGATCTTGCTG 1450
DB 461 PheLeuArgIleValGlnValAsnLeuIleAspPheIleValCysAsnAspTyrLeuVal 480
QY 1451 TATAGACGTTACCTTACCCCAAGAGATGATGTGTGTGGGACCTTCCGGGGGACAGAGAC 1510
DB 481 TyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeuAlaGlyIleArgAsp 500
QY 1511 TCCTGCGCAGGAGACAGCGGGGGGCTTGTCTGTGTGAGCAACACCGCTGTGACCTG 1570
DB 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTyrTyrLeu 520
QY 1571 GCAGGTGTACCAAGCTGGGGGACAGGCTGTGTGGCCAGAGAAACAACCTGTGTGTACCC 1630
DB 521 AlaGlyValThrSerTyrGlyThrGlyCysGlyGlnArgAsnIleProGlyValIleThr 540
QY 1631 AAAGTGAAGAGATCTTCCCTGATTTTACAGCAAGATGAGAGGAGGTGCGA 1684
DB 541 LysValThrGluValLeuProTyrIleTyrSerIleMetGluAsnArgAlaGln 558

```

RESULT 8

AD110395 ID AD110395 standard; protein; 688 AA.

AD110395; AC

22-APR-2004 (first entry)

Human cell surface protease #13.

therapeutic agent; plasmin; protease specific antigen; PSA;
cell-surface protease-associated disease; cancer; ocular disease;
cardiovascular disease; chronic inflammatory disease; wound;
circulatory disorder; dermatological disorder; rheumatoid arthritis;
psoriasis; diabetic retinopathy; pterygium;
excimer laser surgery scarring; glaucoma filtering surgery scarring;
macular degeneration; cret syndrome; solid neoplasm; vascular tumour;
melanoma; Kaposi's sarcoma; human; cell surface protease.

Homo sapiens.

OS

```

PN W0200295007-A2.
XX
PD 28-NOV-2002.
XX
PF 23-MAY-2002; 2002WO-US016819.
XX
PR 23-MAY-2001; 2001US-0293267P.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Madison EL, Sempke JE, Vlausk GP, Kemp SJ, Komandla M, Siev DV;
XX
DR WPI; 2003-221280/21.
DR N-PSDB; AD110394.
XX
PT Novel conjugate useful for treating cell-surface protease-associated
PT disease, comprises a therapeutic agent and a peptidic or nucleic acid
PT substrate linked to it optionally by a peptidic linker.
XX
PS Claim 9; SEQ ID NO 26; 581bp; English.
XX
CC The invention comprises a conjugate that consists of a therapeutic agent
CC and a peptide substrate (optionally linked via linker). The peptide
CC substrate is proteolytically cleaved by a cell surface protease pr a
CC soluble, released or shed form of it, to liberate the therapeutic agent,
CC the conjugate of the invention is not substantially cleaved by plasmin or
CC protease specific antigen (PSA). The conjugate of the invention is useful
CC for treating a cell-surface protease-associated disease such as: cancer,
CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
CC wound, circulatory disorders, dermatological disorders, rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
CC scarring from excimer laser surgery, scarring from glaucoma filtering
CC surgery, macular degeneration, cret syndromes, solid neoplasms, vascular
CC tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence
CC represents a human cell surface protease.
XX
SQ Sequence 688 AA:
XX
Alignment Scores:
Pred. No.: 1,67e-178 Length: 688
Score: 2963.00 Matches: 554
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 88.9% Indels: 0
DB: Gaps: 0
US-10-806-370-11 (1-1748) x AD110395 (1-688)
QY 11 ATGAGAGAGGACAGCGGGAATGATCTCCAGCAAGAACACCTTGAGTGCATCT 70
DB 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAAGCCAGGATCTCCAGCTGGGACACTCCAGCCCGGGCATCTCCAGCCAGGATCT 130
DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAAGCCAGGATCTCCAGCTGGGACACTCCAGCCCGGGCATCTCCAGCCAGGATCT 190
DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAAGCTGTACACTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 250
DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAAGCCAGGATCTCCAGCTGGGACACTTTCAGGTCCTTCATCCGAGGATCTCA 310
DB 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
QY 311 TCCGCGAGGCTGACGCTGGGTGAGCAACCTCCCAACAGAGTACCTGTTAGGACACA 370
DB 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValIleValAlaGlyAlaThr 120
QY 371 CCAAGTGGGGGCTGTACCATCCGATCATCTCTGCGAGTCAGACCAAGCAAGGAGGCC 430

```

```

Db      |||ProValGIyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
Qy      431 ACCGGGGAGAGCCCGAGTACGAGCTGCGCAAGTTCACCTGGCGGAGAGCCAGAAAGAG 490
Db      141 ThrArgIuSerProGIyThrSerIeuProIuSerPheTrpArgGIyGIyGln 160
Qy      491 CTACCGCTCATCGGGTGGGTGCTCTCTCATTTGCCCTGTGTGTTCCTCATCATCTCTC 550
Db      161 LeuProIeuIleGIyCyValIeuIeuIeuIeAlaIeuValIeSerIeuIleIleIeu 180
Qy      551 TTCCAGTTCTGGCAGGGCCACAGAGGATCAGGTACAGAGCAGAGAGGAGAGTCTGCC 610
Db      181 PheGIuPheTrpGIyGlnIyIleThrGIyIleArgTrpIyGlnIuArgGIuSerCySer 200
Qy      611 AAGACGCTGTTCGCTGAGCAGGGGTGGTGGATCTGCAAGCTGAGAGAGAGAGAGTGGG 670
Db      201 LysHlValAlaValArgCySerProIyValAlaProCySerLysLysSerProIuIeuGIy 220
Qy      671 TGGGTGAGGTTCGAGTGGAGCAAGTCTCTGCTTAAATCTACTGTGGTCTCTCCATGAG 730
Db      221 CyValAlaArgPheArgPheTrpArgPheSerIeuIeuIySileTrpSerGIySerSerHl 240
Qy      731 TGGCTTCCCATCTGTGAGCAGAACTGGAATGACTCTCTACTAGAGAGAACTTCCAGAG 790
Db      241 TrpIeuProIleCySerSerSerIeuTrpIeuArgSerTrpSerGIyLysThrCySerGln 260
Qy      791 CTGGGTTTCGAGAGGCTCACCGGCAACGAGGTGGCCACAGGAGATTTGGCAACAGC 850
Db      261 LeuGIyPheGIuSerAlaHlIeArgThrTrpGIuValAlaHlArgArgPheHlIeAsnSer 280
Qy      851 TTCTCAATCTTGAGATACAACTCCACATCCAGAGAAAGCTCCACAGTCTGAATGCCCT 910
Db      281 PheSerIleIeuArgTrpIeAsnSerThrIleGlnIuSerIeuHlIeArgSerGIuCySer 300
Qy      911 TCCACGCGGTATATTCCTCCATGTTCCACTGCGGACTGAGGGCCATGACCGAGCGG 970
Db      301 SerGIuArgTrpIleSerIeuGlnIuCySerHlIeArgLysLysLysLysLysLys 320
Qy      971 ATCGTGGAGGGGCGCTGGCCCTCGGATGAGCAAGTGGCTTGGCAAGTGTGCACTTC 1030
Db      321 IleValGIyGIyAlaIeuAlaSerIeAsnSerIyTrpProTrpGIuValIeSerIeuHlIe 340
Qy      1031 GGACACCAACCACTGTGTGAGAGCAGCTCATTTGACGCCAGAGGTGGTCTCACTGCCGC 1090
Db      341 GIyThrTrpHlIeIleCySerGIyGIyThrIeuIleAsnIleGlnTrpValIeuThrAlaIe 360
Qy      1091 CACTGCTTCTTGTGACCCCGGAGAAAGTCTGAGAGGCTGGAAGTGTACGGGGGACCC 1150
Db      361 HlIeCySerPheIeuValThrArgGIyLysValIeuGIuGIyTrpLysValIyAlaGIyThr 380
Qy      1151 AGCAACCTGCACAGTTCGCTGAGGAGAGCTCCATTCGCGAGATCATATCAACAGCAT 1210
Db      381 SerIeuIeuHlIeGIyLysIeuProGIuAlaIeSerIleAlaGlnIleIleIeAsnSerIeu 400
Qy      1211 TACACCGATGAGAGAGCACTATGACATCGCCCTCATGCGGCTGTCCAAAGCCCTGACC 1270
Db      401 TyrThrArgProIuGIuIeAsnArgPheTrpIeAlaIeIeuMetAlaIeuSerIyProIeuThr 420
Qy      1271 CTGTCCGCTCACTTCACCTGCTTGCCTCCCATGATGACAGACCTTTAGCCTCAAT 1330
Db      421 LeuSerAlaHlIeIleHlIeProAlaIeLysIeuProMetHlIeGIyGlnThrPheSerIeuAsn 440
Qy      1331 GAGACCTGCTGATACAGAGCTTGGCAAGACCGAGGAGAGAGATGAAGAACAATCCCCC 1390
Db      441 GIuHlIeCySerTrpIleThrGIyPheGIyLysThrArgGIuTrpIeAsnArgPheLysThrSerTrp 460
Qy      1391 TTCTCCCGGAGGTGACAGTCAATCTCACTGACATTCAGAAATGCAATGACTTACTGCTC 1450
Db      461 PheIeuArgGIuValGIuValAlaIeuIleAsnIeuIeAsnPheLysLysCySerAsnArgPheLysVal 480
Qy      1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGCTTGTGGGGGAGAGAGC 1510

```

```

Db      481 TyrAsnSerTrpIeIeuThrProArgMetMetCySerAlaGIyAsnIeuArgGIyArgAsp 500
Qy      1511 TCCGTCCAGGGAGACAGCGGGGGGCGCTTGTCTGTGAGCAGAAACCGCTGGTACCTG 1570
Db      501 SerCySerGlnGIyAsnSerGIyGIyProIeuValCySerGIuGlnIeuAsnIeuArgTrpIyIeu 520
Qy      1571 GCAGGTGTACACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAACTGTGTGTACACC 1630
Db      521 AlGIyValThrSerTrpGIyThrGIyCySerGIyGlnIuArgAsnLysProGIyValIyThr 540
Qy      1631 AAAGTGACAGAAAGTTCTTCTCGATTTTACAGCAAGATGAGAGAGAGAGTCCGA 1684
Db      541 LysValThrGIuValIeuProTrpIleTrpIySerLysMetGIuAsnArgAlaGln 558

RESULT 9
ADJ46919
ID      ADJ46919 standard; protein; 688 AA.
AC      ADJ46919;
XX      06-MAY-2004 (first entry)
DT      XX
DE      Human transmembrane serine protease (MTSP)-related polypeptide #3.
KW      Human; transmembrane serine protease; MTSP; cell surface protease;
KW      plasmin; prostate specific antigen; PSA; proliferative disease;
KW      cell-surface protease-associated disease; autoimmune disease;
KW      inflammatory disease; infectious disease; endocrine disease; cancer;
KW      wound; circulatory disorder; dermatological disorder; restenosis;
KW      rheumatoid arthritis; psoriasis; diabetic retinopathy;
KW      laser surgery scarring; glaucoma filtering surgery scarring;
KW      macular degeneration; CRBSF syndrome; bacterial infection; viral disease;
KW      solid neoplasm; vascular tumour; lung; colon; prostate; melanoma;
KW      Kaposi's sarcoma; enzyme.
XX      Homo sapiens.
OS      US2004001801-A1.
PN      01-JAN-2004.
PD      23-MAY-2002; 2002US-00156214.
PF      23-MAY-2002; 2002US-00156214.
PR      23-MAY-2002; 2002US-00156214.
PX      (CORV-) CORVAS INT INC.
PA      Madison IL, Semple JE, Vlasuk GP, Kemp SJ, Komandla M, Slev DV;
PI      WPI; 2004-190126/18.
DR      N-PsDB; ADJ46918.
XX      Conjugate useful for treating e.g. cancer, cell-surface protease-
PT      associated diseases, comprising a peptidic substrate or nucleic acid
PT      substrate linked to a therapeutic agent through a linker.
XX      Claim 9; SEQ ID NO 26; 361pp; English.
XX
XX      The invention relates to a conjugate comprising a therapeutic agent and a
XX      peptidic substrate or nucleic acid substrate linked to the agent
XX      optionally through a linker or peptidic linker, where the peptidic
XX      substrate is proteolytically cleaved by a cell surface protease or a
XX      soluble, released or shed form conjugate to liberate the agent and the
XX      conjugate is not substantially cleaved by plasmin or prostate specific
XX      antigen (PSA). The conjugate is useful for treating a disease, which
XX      involves administering a conjugate to a subject, where the disease is
XX      preferably a proliferative disease or a cell-surface protease-associated
XX      disease. The diseases include autoimmune diseases, inflammatory diseases,
XX      infectious diseases and endocrine diseases. The conjugate is useful for
XX      treating a cell-surface protease-associated disease, which involves
XX      administering a conjugate comprising an agent and a peptidic substrate to
XX      a subject exhibiting symptoms of a cell-surface protease-associated

```

CC disorder, where the disease is selected from cancer, ocular disorders,
 CC cardiovascular disorders, chronic inflammatory diseases, wounds,
 CC circulatory disorders, dermatological disorders, resenosis, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, scarring from laser
 CC surgery, scarring from glaucoma filtering surgery, macular degeneration,
 CC CRIST syndrome, bacterial infections, viral diseases, solid neoplasms and
 CC vascular tumors such as lung cancer, colon cancer, prostate cancer,
 CC melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
 CC serine protease (MSP)-related polypeptide of the invention.

XX Sequence 688 AA:

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 1.67e-178 | Length: | 688 |
| Score: | 2963.00 | Matches: | 554 |
| Percent Similarity: | 99.6% | Conservative: | 2 |
| Best Local Similarity: | 99.3% | Mismatches: | 0 |
| Query Match: | 88.9% | Indels: | 0 |
| DB: | 8 | Gaps: | 0 |

US-10-806-370-11 (1-1748) x ADJ46919 (1-688)

QY 11 ATGGAGAGGAGACACCGGAGATGATCTCCAGCAAGAACCTTACGTGAGCATCT 70
 DB 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAGCCAGGATCTCCAGCTGGAGACACTCCAGGCGGGGATCTCCAGCCAGGATCT 130
 DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAGCCAGGATCTCCAGCTGGAGACACTCCAGGCGGGGATCTCCAGCCAGGATCT 190
 DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAGCTGGATACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 250
 DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAGCCGCGGATCTCCGCTTCTGGCATCACTTTCAGGCTCTCAATCCGAGGATCACT 310
 DB 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
 QY 311 TCCGCGGAGGATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 370
 DB 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
 QY 371 CCAGTGGGGGCTGACCATCCATCATCTCTGCGAGTGACAGCACAGAACAGGCGC 430
 DB 121 ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGCCAGGATGACAGCTGCGCAAGTTTCACTGCGGGAGGCGCAAGAGAG 490
 DB 141 ThrArgGlnSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlnGlnSerGln 160
 QY 491 CTACCGGATCACTGGGCTGCTCTCTCTCAATGCGGCTGGGTTTGGCTATCATCTCTC 550
 DB 161 LeuProLeuIleGlyCysValLeuLeuIleAlaLeuValValSerLeuIleLeuLeu 180
 QY 551 TTCAGGTTCTGGAGGCGCACAGGATGAGTACAGAGAGAGAGAGAGAGAGTGTCTCC 610
 DB 181 PheGlnPheTrpGlnGlyHisThrGlyIleArgGlyLysGlnGlnArgGlnSerCysPro 200
 QY 611 AAGCAGCTGTTCCGCTGACCGGGGTGGTGAATGCAAGTGAAGTGAAGTGAAGTGGC 670
 DB 201 LysHisAlaValArgCysAspGlyValValAlaSerCysLysLeuLysSerAspGlyLeuGly 220
 QY 671 TGGGTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 730
 DB 221 CysValAlaArgPheAspTrpAspLysSerLeuLeuLysLysLysSerGlySerSerHisGln 240
 QY 731 TGGCTTCCCATCTGAGAGAGAACTGAGATGACTCTTACTGAGAGAGAACTGCGCAGAG 790
 DB 241 TrpLeuProIleCysSerSerSerAsnTrpAsnAspSerTyrSerGlnLysThrCysGlnGln 260

QY 791 CTGGGTTTGAGAGTGTCTACCGGACACAGGAGTTGCCACAGGATTTTGGCAAGC 850
 DB 261 LeuGlyPheGlnSerAlaHisAlaGlyThrGlnValAlaHisAlaArgAspPheAlaSer 280
 QY 851 TTCATCACTTGAATCAACTCCACCATCCAGAAAGCTCCACAGGTTGAATGCCCT 910
 DB 281 PheSerIleLeuArgTrpAsnSerThrIleGlnGlnSerLeuHisAlaSerGlnCysPro 300
 QY 911 TCCGAGCGGTATATCTCTCCAGTGTTCCTCCAGTGTTCCTCCAGTGTTCCTCCAGTGTTC 970
 DB 301 SerGlnArgTrpIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
 QY 971 ATCGTGGAGGAGGCGGCTCGGATGAGCAAGGAGGCTTGGCAAGTGAAGTCACTTC 1030
 DB 321 IleValIleValAlaLeuAlaSerSerLysTrpProTrpGlnValSerLeuHisPhe 340
 QY 1031 GGCACCAACCAATCTGTGAGGACCGCTGATGACCCAGTGGTGTCTACTGCGGCC 1090
 DB 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAla 360
 QY 1091 CACTGCTTCTTCTGACCCGGAGAGAGTCTCTGAGGCTGGAAAGTGTACCGCGCAC 1150
 DB 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
 QY 1151 AGCAACCTGACACAGTTCGCGAGGAGCGCTCCATTCGCGAGATCATCATCAAGCAAT 1210
 DB 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleAsnSerAsn 400
 QY 1211 TACACGATGAG 1270
 DB 401 TyrThrAspGlnGlnLysAspArgLysPheIleAlaLeuMetArgLeuSerLysProLeuThr 420
 QY 1271 CTGTCGCTCAATCAACCTCTGCTTCCCTCCCATGATGACAGACCTTTAGCTTCAAT 1330
 DB 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
 QY 1331 GAGACCTGCGAGATCAAGAGCTTTGGCAAGACAGGAGAGAGAGAGAGAGAGAGAGAGAG 1390
 DB 441 GlnThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
 QY 1391 TTCCTCCGAGAGGAG 1450
 DB 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
 QY 1451 TATGACATTAATCTTACCCCAAGATGATGTGTCTGGGACCTTCTGGGGGAGAGAGAC 1510
 DB 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuArgGlyArgAsp 500
 QY 1511 TCTGCGAGGAGAGACAGCGGGGCGCTCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1570
 DB 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
 QY 1571 GCAGGTGTACACAGCTGGGAG 1630
 DB 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
 QY 1631 AAAGTGAAGAAGTTCTTCCCTGATTTTACAGAAATGAGAGAGAGAGAGAGAGAGAG 1684
 DB 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetGlnAsnArgAlaGln 558

RESULT 10
 ADY50054
 ID ADY50054 standard; protein; 688 AA.
 AC ADY50054;
 XX 19-MAY-2005 (first entry)
 DE Human endothelase-2L.
 XX Antiarthritic; Antirheumatic; Antiosteoporotic; Antidiabetic;
 KW Ophthalmological; Cardiovascular-Gen.; Vulnary; Antiinflammatory;

KM Vasotropic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
 KM rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
 KM dermatological disease; cancer; neoplasm; endothelase-2; enzyme.
 OS Homo sapiens.
 XX WO2005019270-A2.
 PN 03-MAR-2005.
 PD 12-AUG-2004; 2004WO-US026148.
 PF 14-AUG-2003; 2003US-0495005P.
 PR 14-NOV-2003; 2003US-0520164P.
 XX (DYAX-) DYAX CORP.
 PA (DENND-) DENNDREON CORP.
 XX Madison EL, Nixon A;
 FI WPI; 2005-202609/21.
 DR N-PSDB; ADY50053.
 PT Novel protein capable of inhibiting endothelase-2, useful for treating
 PT or preventing angiogenesis related disorder e.g. cancer.
 XX Disclosure; SEQ ID NO 2; 157bp; English.
 XX The invention relates to an isolated protein (I) which comprises a heavy
 CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
 CC immunoglobulin variable domain sequence, where the first and second
 CC immunoglobulin variable domain sequences from an antigen binding site
 CC that specifically binds to human endothelase-2. (I) is useful for
 CC detecting an endothelase or endothelase activity in a sample, for
 CC modulating an activity of an ET2-expressing cell, for modulating
 CC proteolysis, for killing or inhibiting growth of a cell, for detecting
 CC endothelase in a subject, for modulating endothelase activity in a
 CC subject, for treating or preventing a disorder characterized by unwanted
 CC angiogenesis in a subject. The disorder is chosen from rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, ocular disorder such as
 CC pterygii recurrence, scarring excimer laser surgery and glaucoma
 CC filtering surgery, cardiovascular disorders, chronic inflammatory
 CC disorders, wound repair, circulatory disorders, crest syndromes,
 CC dermatological disorders and cancers. The present sequence represents the
 CC human endothelase-2L.
 CC
 CC
 SQ Sequence 688 AA;
 Alignment Scores:
 Pred. NO.: 1.67e-178 Length: 688
 Score: 2963.00 Matches: 554
 Percent Similarity: 99.6% Conservative: 2
 Best Local Similarity: 99.3% Mismatches: 2
 Query Match: 88.9% Indels: 0
 DB: 9 Gaps: 0
 US-10-806-370-11 (1-1748) x ADY50054 (1-688)
 QY 11 ATGGAGAGGAGCAGCAGGAGATGATCTTCAGCAAGAACCTTGAGTGAGCATCT 70
 Db 1 MetcunargabserhlsaglyshmlaserProalaagthrrpseralaglyhlsaser 20
 QY 71 CCAGCCCAAGGATCTTCAGCTGGAGACCTTCAGGCCGGGATCTCCAGCCAGGATCT 130
 Db 21 ProalaglnalaserProalaglyThrrProProgllyargalaserProalaglnalaser 40
 QY 131 CCAGCCCAAGGATCTTCAGCTGGAGACCTTCAGGCCGGGATCTCCAGCCAGGATCT 190
 Db 41 ProalaglnalaserProalaglyThrrProProgllyargalaserProalaglnalaser 60
 QY 191 CCAGCTGGTACACCTTCAGGCCGGGATCTTCAGGCCGGGATCTCCAGCCAGGATCT 250
 Db 61 ProalaglyThrrProProgllyargalaserProalaglnalaserProalaglnalaser 80

QY 251 CCAGCCCGGAGATCTCCGGCTTGAGCATCTTCCAGCTTCATCCGGAGGATCA 310
 Db 81 ProalargalaserProalaleuhalaserleuserlrgsserserGlyhrgsser 100
 QY 311 TCCGCAGAGTACGCTTGAGTACAACTCCCAACAGAGTACTCTTTAGAGCAACA 370
 Db 101 SerlaargsseralaserValThrrSerProThrrargVallyrLeuValargAlathr 120
 QY 371 CCAGTGGGGGCTGACCATCCGATCATCTCTCCAGAGTCCAGGACCAAGAGGGCC 430
 Db 121 ProvalgllyalavalProileargsserProalargsseralalathrrargla 140
 QY 431 ACCAGGAGAGCCCAAGGATACAGCCTGCCAAGTTCACTGGCCGGAGGGCCAGAACG 490
 Db 141 ThrrargluserProgllyThrrSerleuProlyserhethrrpargllygllyleu 160
 QY 491 CTACCGCTCATCGGGTGGCTGCTCTCTCTCATTCGCTGGTGGTTGCTCATCTCTC 550
 Db 161 leuProleuileglyCysvalleuileuilealaleuValalaserleuileleu 180
 QY 551 TTCAGTTCTGGCAGGGCCACAGAGGATCAGTACAAGAGCAGAGGAGAGAGCTGCC 610
 Db 181 PheglmphehrrpglnglyhsthrGlyllehrglyrlysglulngluserCysePro 200
 QY 611 AAGCAGCCTGTTCCGCTGACGAGGAGTGGTGGTGACTGCAAGCTGAAGATGACAGCTGG 670
 Db 201 LyshtsalavalargCyseargGlyvalalarpCyseleuLysleuLysleuLysleu 220
 QY 671 TGCCTGAGGTTTGAAGCTGGGACAACTCTCTGTTAAATCTACTCTGGTCTCTCCATCAG 730
 Db 221 CysvalargPhealrppalrppalrppalrppalrppalrppalrppalrppalrppalr 240
 QY 731 TGGCTTCCCATCTGTACAGAGATCGAATGATCTCTACTACAGAGAGACCTGCCAGAG 790
 Db 241 TrpleurolleCysserserhenthrrpshpserlyserglulrThrrCyglngln 260
 QY 791 CTGGGTTTGAAGTGTCTACCCGAGCAACCGAGGTTGCCACAGAGATTTTCCAGACG 850
 Db 261 leuolylehgluseralalhsargThrrThrglvalalhsargalrppalalhsar 280
 QY 851 TTCCTCATCTTGAGATCAACTCCACATCCAGGAAGGCTCCCAAGTCTGAATGGCCT 910
 Db 281 PheaserlleuarglyrpsenSerThrrleglngluserleuLysargsserGlyCysePro 300
 QY 911 TCCAGCGGATATCTCCCTCCAGTGTTCACATCGCGAGCTGAGGAGGATGACCCGGCG 970
 Db 301 SerlmglyrlylleserleuGlnCysehlsCysgllyleuargalameThrrGlyhrg 320
 QY 971 ATCGTGGAGGGGGGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTGGCACTTC 1030
 Db 321 lIvalolyllylaleuhalaserlaserlaserlaserlaserlaserlaserlaserlaser 340
 QY 1031 GGCACCAACCAATCTGTGAGGAGGAGCTGATGACCCCAAGTGGGGTGCACCTGGCC 1090
 Db 341 GlyThrrThrrhlsleCyseGlyGlyThrrleuilealrppalrppalrppalrppalr 360
 QY 1091 CACTGCTTCTTGGTGAACCGGAGAGAGGCTCTGAGAGGCTGGAAGGCTGACGCGAGCC 1150
 Db 361 HlsCysePheheValThrrargglulrvalleuGlnlylrrpysvallyrThralaglyThrr 380
 QY 1211 TACACCGATAGAGAGAGCAGTATGACATGGCCCTCATGGGGCTGTCCAGCCCTGACC 1270
 Db 401 TyThrrpoglulnglulapalrppalrppalrppalrppalrppalrppalrppalrppalr 420
 QY 1271 CTGTCCGCTCACATCCACCTGCTGGCTGCCATGATGACAGACCTTAGGCTCAAT 1330
 Db 421 leuserlathstlehlserProalalCyseleuPromethlsGlyGlnThrrPheaserleuabn 440

| | | | |
|----|------|--|------|
| QY | 431 | ACCAAGAGAGAGCCGAGTATACAGAGCTGGCCAAAGTTCACTGGCGGAGAGCCAGAAAGAG | 490 |
| Db | 141 | ThraTgJusSerProGlyThrSerLeuProLysPheThrTTPAArgGluGlyGlnSerGln | 160 |
| QY | 491 | CTACCGCTCATGGAGGTGGGTGGTCTCTCCCTCAATTGGCCGTGGTGGTTCCCTCATCATCTCTC | 550 |
| Db | 161 | LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIleLeu | 180 |
| QY | 551 | TTCCAGTTCTGGCAGGGGCCACACAGAGATCAGTATCAAGAGACAGAGGAGAGCTGTGCC | 610 |
| Db | 181 | PheGlnPheTrpGlnGlyNH1SerThrGlyIleArgNH1SerGlnArgGlnArgGlySerCysPro | 200 |
| QY | 611 | AAGCAGCGGTCTCGCTGTGACCGGGGTGGTGGATCTGCAAGCTGAAGAGTGAAGAGTGGGC | 670 |
| Db | 201 | LysNH1SerAlaValArgCysAspSerGlyValValAspCysValLeuLysSerAspGluLeuGly | 220 |
| QY | 671 | TGCGTAGAGTTGACTGGGACACAAAGTCTGTGTTAAATCTACTCGGTCTCCCATAG | 730 |
| Db | 221 | CysValArgPheAspTrpAspLysSerLeuLeuValIleTyrSerGlySerSerNH1SerGln | 240 |
| QY | 731 | TGGCTTCCCATCTGTAGACAGCAACTGGAAATGATCTCTACTGACAGAAAGACCTTGGCCAGAG | 790 |
| Db | 241 | TrpLeuProIleCysSerSerSerNH1TrpAsnAspSerTyrSerGluLysThrCysGlnGln | 260 |
| QY | 791 | CTGGGTTTCGAGAGTGCATCCCGGACAACCGAGTGGCCACAGGATTTTGGCCAAAGC | 850 |
| Db | 261 | LeuGlyTrpGlnSerAlaNH1SerThrThrGluValAlaNH1SerAspPheAlaAsnSer | 280 |
| QY | 851 | TTCTCAATCTTGAGATATACACTCCACACATCCAGAAAGCTTCCACAGGTCTGAATGCCCT | 910 |
| Db | 281 | PheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuNH1SerSerGluCysPro | 300 |
| QY | 911 | TCCAGCGGTATATCTCCCTCCAGTGTTCACATGCGGACTGAGGCGCATGACCGGCGG | 970 |
| Db | 301 | SerGlnArgTyrIleSerLeuGlnCysSerNH1CysGlyLeuArgAlaMetThrGlyArg | 320 |
| QY | 971 | ATCGTGGAGAGGGGCGCTGGCTCTGGATATGACAGTGGCTTGGCAAGTGAAGTCTGCATTC | 1030 |
| Db | 321 | IleValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuNH1SerPhe | 340 |
| QY | 1031 | GGCACACCCACATCTGTGAGAGGACGCGCTATGACGCGCGAGTGGGTGCTCACTGGCGCC | 1090 |
| Db | 341 | GlyThrThrNH1SerCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla | 360 |
| QY | 1091 | CAGTGCCTTCTTGATACCCCGGAGAGAGTCTGTGAGGGCTGGAAGTGTACGCGGACCC | 1150 |
| Db | 361 | HisCysPhePheValThrArgGluLysValLeuGlnGlyTrpLysValTyrAlaGlyThr | 380 |
| QY | 1151 | AGCAACCTGCACACAGTTGCTTGAAGGACGCTTCATTTGCCAGATATATCAACAGCAAT | 1210 |
| Db | 381 | SerAsnLeuNH1SerGlnLeuProGlnAlaLeuSerIleAlaGluIleIleIleAsnSerAsn | 400 |
| QY | 1211 | TACACCGATGAGAGAGACATATGACATGCGCCCATGGCGGCTGTGCCAACCCCTGAC | 1270 |
| Db | 401 | TyrThrAspGlnGluAspAspTyrAspIleAlaLeuMetValGluSerLysProLeuThr | 420 |
| QY | 1271 | CTGTCCGCTCATCCACCCCTGTCTTGCCTCCCATGATGACAGACCTTATAGCTCAAT | 1330 |
| Db | 421 | LeuSerAlaNH1SerIleHisProAlaCysLeuProMetNH1SerGlnThrPheSerLeuAsn | 440 |
| QY | 1331 | GAGACTGCTGGATATACAGGCTTTTGGCAAGACGAGAGACATGACATGACATATCCCCC | 1390 |
| Db | 441 | GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysThrSerPro | 460 |
| QY | 1391 | TTCCCTCGGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTGGTC | 1450 |
| Db | 461 | PheLeuArgGluValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeuVal | 480 |
| QY | 1451 | TATGACAGTTACTTAAACCCCAAGATGATGTGTGTGGAGACCTTCGAGGGGACAGAC | 1510 |
| Db | 481 | TyrAspSerTyrLeuThrThrProArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp | 500 |
| QY | 1511 | TCCTGCGAGGAGACAGCGGGGGGCTCTTGTCTGTGACAGAAACAACCGCTGGATCACTG | 1570 |

| | | | | |
|------------------|---|------|---|------|
| Dd | | 501 | SetoyseqlngllyapseserqlygilyproleuvaLcyoeluglnmenamargtpryleu | 520 |
| Oy | | 1571 | GCAAGTGTACCCAGCTGGGACACAGCGTGTCGCCAGAAACAACCTGGTGTATCACC | 1630 |
| Dd | | 521 | AlaelylalthrserTPglYthnQlyCyegilYcInlrgrmshysProglYValTYThr | 540 |
| Oy | | 1631 | AAAATGCACAGAATTCTTCCTCGATTATACGCAAGATGGAG | 1672 |
| Dd | | 541 | LysValThrgluValLeuProtIptIetyserLyewetglu | 554 |
| RESULT 12 | | | | |
| ID | ADH17430 | | | |
| XX | ADH17430 standard; protein: 593 AA. | | | |
| XX | ADH17430; | | | |
| XX | AC | | | |
| XX | DT | | | |
| DE | 11-MAR-2004 (first entry) | | | |
| XX | | | | |
| DE | Human NOVI21 protein - SEQ ID 120. | | | |
| XX | | | | |
| KW | NOVI; antidiabetic; anorectic; cardiact; hypotensive; | | | |
| KW | antihistaminic; anorectic; virtucide; antibacterial; fungicide; | | | |
| KW | protosaccharide; noctropic; neuroprotective; antiapoptosis; | | | |
| KM | anticonvulsant; osteopathic; antiarthritic; antiinflammatory; | | | |
| KM | dematological; antiaesthetic; antilipemic; metabolic; diabetes; | | | |
| KW | obesity; infection; anorexia; cancer; cardiovascular; hypertension; | | | |
| KW | atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's; | | | |
| KW | epilepsy; immune; osteoarthritis; haemopoietic; | | | |
| KW | inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis; | | | |
| KW | cell differentiation; proliferation; haemopoiesis; wound healing; | | | |
| KW | angiogenesis; gene therapy; chromosome mapping; tissue typing; | | | |
| KW | pharmacogenomic; human. | | | |
| XX | | | | |
| OS | Homo sapiens. | | | |
| XX | | | | |
| PN | WO2003093432-A2. | | | |
| XX | | | | |
| PD | 13-NOV-2003. | | | |
| XX | | | | |
| PF | 02-MAY-2003; 2003WO-US013690. | | | |
| XX | | | | |
| PR | 02-MAY-2002; 2002US-0377321P. | | | |
| PR | 08-MAY-2002; 2002US-0378730P. | | | |
| PR | 24-MAY-2002; 2002US-0383075P. | | | |
| PR | 29-MAY-2002; 2002US-0384044P. | | | |
| PR | 30-MAY-2002; 2002US-0384215P. | | | |
| PR | 30-MAY-2002; 2002US-0384286P. | | | |
| PR | 30-MAY-2002; 2002US-0384287P. | | | |
| PR | 30-MAY-2002; 2002US-0384337P. | | | |
| PR | 30-MAY-2002; 2002US-0384352P. | | | |
| PR | 31-MAY-2002; 2002US-0385211P. | | | |
| PR | 02-JUL-2002; 2002US-0393333P. | | | |
| PR | 09-AUG-2002; 2002US-0402154P. | | | |
| PR | 09-AUG-2002; 2002US-0402171P. | | | |
| PR | 09-AUG-2002; 2002US-0402204P. | | | |
| PR | 09-AUG-2002; 2002US-0402205P. | | | |
| PR | 22-AUG-2002; 2002US-0405175P. | | | |
| PR | 27-AUG-2002; 2002US-0406129P. | | | |
| PR | 27-SEP-2002; 2002US-0412954P. | | | |
| PR | 30-SEP-2002; 2002US-0414975P. | | | |
| PR | 07-OCT-2002; 2002US-0416611P. | | | |
| PR | 24-OCT-2002; 2002US-0420851P. | | | |
| PR | 31-OCT-2002; 2002US-0422547P. | | | |
| PR | 01-MAY-2003; 2003US-00428275. | | | |
| XX | | | | |
| PA | (CURA-) CURAGEN CORP. | | | |
| XX | | | | |
| PI | Alvarez ER, Anderson DW, Boldog FT, Catterton E, Edinger SR, | | | |
| PI | Fernandes BR, Gerlach VL, Gorman L, Grose WM, Guo X, Ji W, | | | |
| PI | Kekuda R, Li L, MacDougall JR, Padigar M, Patturajan M, | | | |
| PI | Peterzon JM, Rastelli L, Shiwets RA, Spytek KA, Stone DJ, | | | |
| PI | Vernet CM, Voss EZ, Zhong M; | | | |

XX WPI: 2004-053040/05.
DR N-PSDB: ADH17429.

PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

PS Claim 1; SEQ ID NO 120; 478bp; English.

XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiatic, hypotensive, antiarteriosclerotic, anorectic, vitricide,
CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteoprotective, antianthratic,
CC antiinflammatory, dermatological, antisthmatic and antileptic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis. In gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.

XX Sequence 593 AA:

Alignment Scores:

Pred. No.: 6,49e-178 Length: 593
Score: 2953.50 Matches: 555
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 5
Query Match: 88.6% Indels: 0
DB: 8 Gaps: 1

US-10-806-370-11 (1-1748) x ADH17430 (1-593)

QY 8 ACATGAGAGAGGACAGCCAGGAGATGATCTCCAGCAGAGACCTTCAGTGGAGCA 67
DB 4 ThiretGUAAGABSRHISGLYASnhlaSerProhlaArgThnProserAlaGlyAla 23
QY 68 TCTCCAGCCAGGAGATCTCCAGCTGGAGACACTTCAGGCGGAGCATCTCCAGCCAGGCA 127
DB 24 SerProhlaGlnAlaSerProhlaGlyThnProhlaGlyArgAlaSerProhlaGlnAla 43
QY 128 TCTCCAGCCAGGAGATCTCCAGCTGGAGACACTTCAGGCGGAGCATCTCCAGCCAGGCA 187
DB 44 SerProhlaGlnAlaSerProhlaGlyThnProhlaGlyArgAlaSerProhlaGlnAla 63
QY 188 TCTCCAGCTGGATACCTCCAGGCGGAGCATCTCCAGGCGG-----GCA 232
DB 64 SerProhlaGlyThnProhlaGlyArgAlaSerProhlaGlyArgAlaSerProhlaGlnAla 83
QY 233 TCTCCAGCCAGGAGATCTCCAGGCGGAGCATCTCCAGCTGGAGATCTTCAGAGTCC 292
DB 84 SerProhlaGlnAlaSerProhlaGlyArgAlaSerProhlaGlnAlaSerLeuSerArgSer 103
QY 293 TCATCCGAGAGTATCATCTCCAGGCGGAGCATCTCCAGCTGGAGATCTTCAGAGTCC 352
DB 104 SerSerGlyArgSerSerSerAlaArgSerAlaSerValThnThnSerProhlaArgVal 123
QY 353 TACCTTTGTAAGCAACACAGTGGGGGCTGTACCATCCAGTATCTCTCCAGAGTCA 412
DB 124 TyrlauValaArgAlaThnProhlaGlyAlaValaProhlaArgSerSerProhlaArgSer 143

QY 413 GCACAGCAACACAGGCGCACAGAGAGAGCCAGGTACGAGCTGCGCAAGTTCACCTGG 472
DB 144 AlaProhlaThnArgAlaThnArgGlySerProhlaGlyThnSerLeuProhlaSerPheThnTrp 163
QY 473 CGGAGAGGCGCAGAGAGAGTACCGCTTCATCGGAGTGGTCTCTCTCATTCGCGCTGGT 532
DB 164 ArgGlyGlyGlnGlySerGlnLeuProhlaGlyGlyCysValLeuLeuLeuLeuAlaLeuVal 183
QY 533 GTTTCGCTCATCATCTCTCCAGTTCCTGAGAGGCGCACAGAGATCAGTCAAGAG 592
DB 184 ValSerLeuLeuLeuLeuPheGlnPheTrpGlnGlyAlaThnTrpGlyLeuGlyGlyGln 203
QY 593 CAGAGGAGAGAGTGTCCCAAGACAGCTGTTCGCTGAGCAGGAGTGGTGAATCAAGCTG 652
DB 204 GlnArgGlySerCysProhlaSerAlaValaArgCysAspGlyAlaValaArgCysVala 223
QY 653 AAGAGTACAGAGTGGAGTGGCTGAGAGTTCAGTGGAGCAAGTCTCTGTTAAATCTAC 712
DB 224 LysSerAspGlyLeuGlyCysValaArgPheAspTrpAspLysSerLeuLeuLysIleTyr 243
QY 713 TCTGGGCTCTCCATCATGAGTGGCTTCATGTGAGACAGCACTGGATGATCTCTACTCA 772
DB 244 SerGlySerSerHisGlnTrpLeuProhlaCysSerSerSerAntTrpAspSerTyrSer 263
QY 773 GAGAGACCTGCGCAGAGAGTGGGTTTCAGAGTGTCAACCGGCAACCGAGTTCGCCAC 832
DB 264 GlnLysThnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 283
QY 833 AGGAGATTTTCGCAACAGCTTCATCTTGAATGATACATCCACATCCAGAGAAAGCTTC 892
DB 284 ArgAspPheAlaAsnSerPheSerIleLeuArgTyrAsnSerThrIleGlnGlnGlnSerLeu 303
QY 893 CACAGCTGTGAATGCCCTTCCACAGCGGATATATCTCTCCAGTGTTCACCTGCGAGCTG 952
DB 304 HisArgSerGlyCysProhlaGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeu 323
QY 953 AGGCGCATGACCGGCGGAGATCTGGAGAGGCGGCTCGGATACAGTGGAGCTTCGG 1012
DB 324 ArgAlaMetThrGlyArgGlyLeuAlaGlyAlaLeuAlaSerAspSerLysTrpProTrp 343
QY 1013 CAAGTGAAGTGTGACATTCGAGCAGCAGCAGCAGATCTGTGAGAGCAGCTCATTCAGCGCCAG 1072
DB 344 GlnValSerLeuHisPheGlyThnThnHisIleCysGlyGlyThnLeuIleAspAlaGln 363
QY 1073 TGGGTCTCACTGCGCGGCCACTGCTTTTGTGTGACCCGAGAGAGCTCTGGAGGCTGG 1132
DB 364 TrpValLeuThnAlaAlaHisCysPhePheValThnArgGlyLysValaLeuGlnGlyTrp 383
QY 1133 AAGGTGATCCGCGGAGCAGCAGCAGCAGTTCGAGGAGCTGAGGAGCTCATTCAGCGGAG 1192
DB 384 LysValaLysAlaGlyThnSerAsnLeuHisGlnLeuProhlaGlnAlaAlaSerIleAlaGln 403
QY 1193 ATCATCATCAACAGCAATTCACCGATGAGAGAGCAGTATGACATCGGCTTCATCGG 1252
DB 404 IleIleIleLeuSerAsnTrpTrpIleAspGlnGlnAspArgLysIleAlaLeuMetArg 423
QY 1253 CTGTCAAGGCGCTGAGCTGTGCGGCTGACATCCACCTGCTGCTGCTCCCATGATGAGA 1312
DB 424 LeuSerLysProhlaLeuThnLeuSerAlaHisIleHisProhlaCysLeuProhlaMetHisGly 443
QY 1313 CAGACCTTTAGAGCTGATGAGACCTGTGATGATGACAGGCTTTGGCAGAGCAGGAGAGA 1372
DB 444 GlnThnPheSerLeuAsnGlnThnCysTrpIleThnIlePheGlyLysThnArgGlnThn 463
QY 1373 GATGACAAGACATCCCTCTCTCCGAGAGTGGAGGTCATCTCATGCAATTCAGAGAA 1432
DB 464 AspAspLysThrSerTrpPheLeuArgGlnValaGlnAlaLeuLeuIleAspPheLysLys 483
QY 1433 TGCAATGATCTTGGTGTATGACAGTTCATCCCAAGATGATGTGTGCTGGGAGC 1492
DB 484 CysAsnAspLysLeuValaLysAspSerTyrLeuThnProhlaArgMetCysAlaGlyAsp 503

Db 121 LeuValAlaAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
 Qy 416 CCGACACCGAGGGGACAGGGAGAGCCCGAGTACGAGCCCTGACGATTCACCTGCGG 475
 Db 141 ProIleAlaThrAlaAlaThrArgIleSerProGlyThrSerLeuProLysPheThrTrpArg 160
 Qy 476 GAGGGCCAGAAAGACCTACCGCTCATCGGGGCTGCTCCCTCATTTGACCTGGTGT 535
 Db 161 GluGlyIleuLysGlnLeuProIleGlyCysValLeuLeuLeuIleAlaLeuVal 180
 Qy 536 TCCTCATCATCTCTTCCAGTTCTGCGAGGGCCACAGGGATCAGTACAGAGCAG 595
 Db 181 SerIleIleIleLeuPheGlnPheThrGlnIleIleThrGlyIleArgGlyLysGln 200
 Qy 596 AGGAGAGCTGTCCCAAGCAGCTGTTCGTGTACCGGGGCTGGAGCTGAACTGAAG 655
 Db 201 ArgGlySerCysProLysAlaValArgCysAspGlyValAlaAspCysLysLeuLys 220
 Qy 656 AGTACAGAGCTGGGCTGGAGGTTGACTGGGACAACTGGAAATGACTCTTACTCAG 715
 Db 221 SerAspIleuLeuGlyCysValArgPheAspTrpAspLysSerLeuLeuLysIleTrpSer 240
 Qy 716 GGGTCTCCCATCATGAGTGGCTTCCCATCTGTAGACAACTGGAAATGACTCTTACTCAG 775
 Db 241 GlySerSerHisGlnTrpLeuProIleCysSerSerIleTrpAsnAspSerTrpSerGln 260
 Qy 776 AAGACCTCCACGACGCTGGGTTTGGAGAGTCTTACCGGAGAACCGAGGTTGCCACAG 835
 Db 261 LysThrCysGlnGlnLeuGlyPheGlnSerAlaHisArgThrTrpGluValAlaHisArg 280
 Qy 836 GATTTTGGCAACAGCTTCTCAATCTTGAGATACAACTCCACATCCAGAAAGCTCCAC 895
 Db 281 AspheAlaAsnSerPheSerIleLeuArgTrpAsnSerThrIleGlnIleSerLeuHis 300
 Qy 896 AGGTGTGAATGCCCTTCCACAGCGGATATATCTCCCTCCAGTGTTCACCTGCGAGCTGAG 955
 Db 301 ArgSerGlnCysProSerGlnArgTrpIleSerLeuGlnCysSerHisCysGlyLeuArg 320
 Qy 956 GCCATGACCGGGGCGGATCTGTGGAGGGGCGCTGGCTGGATACAGAGTGGCTTGGCAA 1015
 Db 321 AlaMetThrGlyAlaGlyIleValGlyAlaLeuAlaSerAspSerLysTrpProTrpGln 340
 Qy 1016 GTGAGTCTGCACTTCGGAGCACCACATCTGTGGAGGCAAGCTTGAAGGCCACAGG 1075
 Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrp 360
 Qy 1076 GTGCTCACTGCGCGCCACTGCTTCTCTGTGACCCGCGAGAAAGTCTGAGAGGCTGAG 1135
 Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgLysValLeuGlnGlyTrpLys 380
 Qy 1136 GTTGTAGCGGGGACCAAGCAACCTGACCAAGTGTGCTGAGGAGCGCTTCATTTGCCGAGATC 1195
 Db 381 ValLysAlaGlyThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIle 400
 Qy 1196 ATCATCAACAGCAATTACACCGATGAGAGAGACACTTATGATCCGCTCATCCGCTG 1255
 Db 401 IleIleAsnSerAlaGlyThrAspGlnGlnAspAspTrpAspIleAlaLeuMetArgLeu 420
 Qy 1256 TCCAAGCCCGCTGAGCTGTCCGCTCACATCCACCTGCTTCCCATGCTCATGAGTGAAG 1315
 Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuPromethHisGlyGln 440
 Qy 1316 ACCTTTAAGCTCAATGAGACCTGTGTGATCAAGAGCTTTGGCAAGACAGGAGACAGAT 1375
 Db 441 ThrPheSerLeuAsnGlnThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAsp 460
 Qy 1376 GACAAGACATCCCTTCCTCCGGAGGCTGAGGCAATCATCGACTTCAAGAAATGC 1435
 Db 461 AspLysThrSerProPheLeuArgIleValAlaAsnLeuIleAspPheLysLysCys 480
 Qy 1436 AATGACTACTTGTGTATGACAGTATACCTTACCCCAAGATGATGTGTGCGGAGACTT 1495
 Db 481 AsnAspTrpLeuValLysAspSerTrpLeuThrProArgMetMetCysAlaGlyAspLeu 500

Qy 1496 CTTGGGGGACAGAGACTCTCTCCAGGAGACAGCGGGGCGCTTGTGTGTGACAGAAC 1555
 Db 501 ArgGlyCylArgArgAspSerCysGlnGlyAspSerCylGlyProLeuValCysGlnGlnAsn 520
 Qy 1556 AACCGCTGTACTGTGCGAGGTTTCACCAAGCTGGGGGACAGGCTGTGGCCAGAGAACAA 1615
 Db 521 AsnArgTrpTrpLeuAlaGlyValThrSerTrpGlyThrGlyCysGlnGlnArgAsnLys 540
 Qy 1616 CTTGTGTGTACACCAAGTACAGAGAGTCTTCCCTGGATTTTACAGCAAGATGAG 1672
 Db 541 ProGlyValLysThrLysValThrGluValLeuProTrpIleTrpSerLysMetGln 559
 RESULT 14
 ADH17434
 ID ADH17434 standard; protein: 586 AA.
 XX
 AC ADH17434;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human NOV12k protein - SEQ ID 124.
 XX
 KW NOX; antidiabetic; anorectic; cardiatic; hypotensive;
 KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
 KW protozoacide; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
 KW dermatological; antiashmatic; antilipemic; metabolic; diabetes;
 KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
 KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
 KW epilepsy; immune; osteoarthritis; haemopoietic;
 KW inflammatory skin disorder; asthma; dyaliidaemia; neurogenesis;
 KW cell differentiation; proliferation; haemopoiesis; wound healing;
 KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
 KW pharmacogenomic; human.
 OS Homo sapiens.
 XX
 PN M02003093432-A2.
 XX
 PD 13-NOV-2003.
 XX
 PF 02-MAY-2003; 2003WC-US013690.
 XX
 PR 02-MAY-2002; 2002US-0377321P.
 PR 08-MAY-2002; 2002US-0378750P.
 PR 24-MAY-2002; 2002US-0383075P.
 PR 29-MAY-2002; 2002US-0384044P.
 PR 30-MAY-2002; 2002US-0384215P.
 PR 30-MAY-2002; 2002US-0384286P.
 PR 30-MAY-2002; 2002US-0384287P.
 PR 30-MAY-2002; 2002US-0384327P.
 PR 30-MAY-2002; 2002US-0384352P.
 PR 31-MAY-2002; 2002US-0385211P.
 PR 02-JUL-2002; 2002US-0393333P.
 PR 09-AUG-2002; 2002US-0402154P.
 PR 09-AUG-2002; 2002US-0402171P.
 PR 09-AUG-2002; 2002US-0402204P.
 PR 09-AUG-2002; 2002US-0402205P.
 PR 22-AUG-2002; 2002US-0405175P.
 PR 27-AUG-2002; 2002US-0406129P.
 PR 23-SEP-2002; 2002US-0412954P.
 PR 30-SEP-2002; 2002US-0414975P.
 PR 07-OCT-2002; 2002US-0416661P.
 PR 24-OCT-2002; 2002US-0420851P.
 PR 31-OCT-2002; 2002US-0422547P.
 PR 01-MAY-2003; 2003US-00428275.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alvarez E, Anderson DW, Boidog FL, Catterton E, Edinger SR,
 PI Fernandes ER, Gerlach VL, Gorman L, Grose WM, Guo X, Ji W,
 PI Kekuda R, Li L, Macdougall JR, Padigaru M, Patunrajan M,

PI Peterson JD, Rastelli L, Shimkets RA, Spytek KA, Stone DJ;
PI Vernet CAM, Voss EZ, Zhong M;
XX MPI: 2004-053040/05.
DR N-PSDB: ADH17433.
XX
PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
XX
PS Claim 1; SEQ ID NO 124; 478bp; English.
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiant, hypotensive, antierosclerotic, anorectic, virucide,
CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antirheumatic,
CC antiinflammatory, dermatological, antiaesthetic and antipneumic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis. In gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.
XX
XX Sequence 586 AA;
SQ
Alignment Scores:
Pred. No.: 1.34e-177 Length: 586
Score: 2948.50 Matches: 554
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 5
Query Match: 88.5% Indels: 5
DB: 8 Gaps: 1
US-10-806-370-11 (1-1748) X ADH17434 (1-586)
QY 11 ATGAGAGGAGGACAGCCACGGGATGATCTTCAGCAAGAACCTTGAGTGAGCATCT 70
Db 1 MerglurrgaspserrHlsglyAanhlaserProhlargrhtrProsserHlsglyAlaser 20
QY 71 CCAAGCCAGGATCTTCAGCTGGGACACCTTCAGAGCCGGGATCTTCAGCCAGGATCT 130
Db 21 ProhlaelmlaserProhlaelgylthtrProhlaelaserProhlaelmlaser 40
QY 131 CCAAGCCAGGATCTTCAGCTGGGACACCTTCAGAGCCGGGATCTTCAGCCAGGATCT 190
Db 41 ProhlaelmlaserProhlaelgylthtrProhlaelaserProhlaelmlaser 60
QY 191 CCAAGCTGTACACTTCAGAGCCGGGATCTTCAGAGCCGG-----GCATCT 235
Db 61 ProhlaelgylthtrProhlaelgylthtrProhlaelaserProhlaelgylthtr 80
QY 236 CCAAGCCAGGATCTTCAGAGCCGGGATCTTCAGAGCCGGGATCTTCAGAGCTCTCA 295
Db 81 ProhlaelmlaserProhlaelgylthtrProhlaelaserProhlaelmlaser 100
QY 296 TCCGAGAGGATCATATCGAGAGGTCAGGCTCGGTGACAACTCCCAACAGAGGTATC 355
Db 101 SerhlyrgerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 120
QY 356 CTGTGTAGAGCAACAGAGTGGGGCTGTACCCATCCGATCATCTCTGACAGGTCA 415

Db 121 LeuValrgrAlatrhrProValGlyAlaValProhlaelrgerSerProhlaelrgerSer 140
QY 416 CCAAGCAACAGAGGCTCCAGAGAGAGCCAGGATACAGAGCTTCAGGATCTTACCTGGCG 475
Db 141 ProhlaelrgrAlatrhrgrAlatrgrlaserProhlaelrgrAlatrgrAlatrgr 160
QY 476 GAGAGCCAGAGAGGATACAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGG 535
Db 161 Glulgllylnygllylnygllylnygllylnygllylnygllylnygllylnygllyln 180
QY 536 TCCGTCATCATCTCTTCAGAGTTCGAGAGGCGCACAGAGGATCAGATACAGAGAGAG 595
Db 181 SerleullelleleupheglInphetrglnclylnshrglylnlearglylnygllyln 200
QY 596 AGGAGAGGCTTCAGAGAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAG 655
Db 201 ArggllyserCyserProhlyshlsAlaValArgCyserProhlyshlsAlaValArg 220
QY 656 AGTACAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCT 715
Db 221 SerhlaeluleuglyCyserValargphaserPthphlyserSerleullellelle 240
QY 716 GGGTCTCCATCATAGTGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCT 775
Db 241 GlyserSerHlsglntPrleupProhlaelCyserSerSerSerSerSerSerSer 260
QY 776 AAGACCTGCAGAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCT 835
Db 261 LysHtrCyserlnglnlnygllylnygllylnygllylnygllylnygllylnygllyln 280
QY 836 GATTTGCCAAGAGCTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 895
Db 281 AsphphelaanaserPheserlleleuHrgrlyrHanserHrHlsgllylnygllyln 300
QY 896 AGGTGTGAATGCCCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAG 955
Db 301 ArgsergllyCyserProserglncArglyrlyleleleuHlncysSerHlsglylnygl 320
QY 956 GCCATGACCGGAGGATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1015
Db 321 AlamerHtrgllyrgrlylnygllylnygllylnygllylnygllylnygllylnygl 340
QY 1016 GTAGAGTGTGACCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAG 1075
Db 341 ValSerleullellellellellellellellellellellellellellellellelle 360
QY 1076 GTGCTCATCTCCGCGCCACTCTTCGTGAGCCCGGAGAGAGGCTTCAGAGGCTTCAG 1135
Db 361 Valleullellellellellellellellellellellellellellellellellelle 380
QY 1136 GTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1195
Db 381 VallyrAlaglylnygllylnygllylnygllylnygllylnygllylnygllylnygl 400
QY 1196 ATCATCAAGAGCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1255
Db 401 lle 420
QY 1256 TCCAGGCGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGG 1315
Db 421 SerlyserProhlyshlsAlaValArgCyserProhlyshlsAlaValArg 440
QY 1316 ACCTTTAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGGCTTCAGAGG 1375
Db 441 ThrPheserleullellellellellellellellellellellellellellellelle 460
QY 1376 GACAAGAGATCCCGCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1435
Db 461 AsphlyrHtrserProhlaelrgrlylnygllylnygllylnygllylnygllylnygl 480
QY 1436 AATGACTACTTGTGTATAGACAGTTACCTTACCCCAAGATGATGTGTGTGGAGCTT 1495

Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeu 500
QY 1496 CGTGGGGCGAGACCTCTCCAGGAGAGAGGGGGGCTCTTGGCTTGACAGAAC 1555
Db 501 ArgGlyGlyArgAspSerCysGlnGlyAspSerGlyAlaProLeuValCysGlnGlnAsn 520
QY 1556 AACCGGTGACTGTCGAGGTGTACACAGCTGGGGGACAGCTGTGGCCAGAGAACAA 1615
Db 521 AsnArgTyrTyrLeuAlaGlyValThrSerTyrGlyThrGlyCysGlyGlnArgAsnLys 540
QY 1616 CTTGTGTGTACACCAAGTTCCTTCCTGATTTACAGCAAGATGAG 1672
Db 541 ProGlyValTyrThrLysValThrGluValLeuProTyrPheTyrSerLysMetGln 559
RESULT 15
ADH17452
ID ADH17452 standard; protein; 586 AA.
XX
XX ADH17452;
XX
XX 11-MAR-2004 (first entry)
DE Human NOV12c protein - SEQ ID 142.
XX
XX NOVX; antidiabetic; anorectic; cardiast; hypotensive;
KM antidiabetic; anorectic; vitruide; antibacterial; fungicide;
KM protozoacide; neuroprotective; antiparkinsonian;
KM anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KM dermatological; antiaesthetic; antidiabetic; metabolic; diabetes;
KM obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KM atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
KM epilepsy; immune; osteoarthritis; haemopoietic;
KM inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
KM cell differentiation; proliferation; haemopoiesis; wound healing;
KM angiogenesis; gene therapy; chromosome mapping; tissue typing;
KM pharmacogenomic; human; SNP; single nucleotide polymorphism.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH MISC-difference 401
FT /note= "Wild-type Thr may be substituted by Ile due to
FT single nucleotide polymorphism (SNP)"
XX
XX MO2003093432-A2.
XX
XX 13-NOV-2003.
XX
XX 02-MAY-2003; 2003WC-US013690.
XX
XX 02-MAY-2002; 2002US-0377321P.
XX 08-MAY-2002; 2002US-0378730P.
XX 24-MAY-2002; 2002US-0383075P.
XX 29-MAY-2002; 2002US-0384044P.
XX 30-MAY-2002; 2002US-0384215P.
XX 30-MAY-2002; 2002US-0384286P.
XX 30-MAY-2002; 2002US-0384297P.
XX 30-MAY-2002; 2002US-0384337P.
XX 30-MAY-2002; 2002US-0384352P.
XX 31-MAY-2002; 2002US-0385211P.
XX 02-JUL-2002; 2002US-0393333P.
XX 09-AUG-2002; 2002US-0402154P.
XX 09-AUG-2002; 2002US-0402171P.
XX 09-AUG-2002; 2002US-0402204P.
XX 09-AUG-2002; 2002US-0402205P.
XX 22-AUG-2002; 2002US-0405175P.
XX 27-AUG-2002; 2002US-0406129P.
XX 23-SEP-2002; 2002US-0412954P.
XX 30-SEP-2002; 2002US-0414975P.
XX 07-OCT-2002; 2002US-0416611P.
XX 24-OCT-2002; 2002US-0420851P.
XX 31-OCT-2002; 2002US-0422547P.
XX 01-MAY-2003; 2003US-00428275.

XX
PA (CURA-) CURAGEN CORP.
XX
PI Alvarez E, Anderson DW, Boldog FI, Catterton E, Edinger SR,
PI Fernandez ER, Gerlach VP, Gorman L, Grosse WM, Guo X, Ji W,
PI Kekuda R, Li L, MacDougall JR, Padigaru M, Patirajan M,
PI Peterson JD, Raselli L, Shimkets RA, Splek KA, Stone DJ,
PI Vernet CM, Voss EZ, Zhong M;
XX
DR WPI; 2004-053040/05.
DR N-PSDB; ADH17451.
XX
XX New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
XX
XX Claim 1; SEQ ID NO 142; 478bp; English.
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiast, hypotensive, antidiabetic, anorectic, vitruide,
CC antibacterial, fungicide, protozoacide, neurotropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic,
CC antiinflammatory, dermatological, antiaesthetic and antidiabetic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.
XX
XX SQ Sequence 586 AA;
SQ
Alignment Scores:
Pred. No.: 2,77e-177 Length: 586
Score: 2943.50 Matches: 553
Percent Similarity: 98.9% Conservative: 0
Best Local Similarity: 98.9% Mismatches: 1
Query Match: 88.3% Indels: 5
DB: Gaps: 1
US-10-806-370-11 (1-1748) x ADH17452 (1-586)
QY 11 ATGGAGAGGGAGCGACCGGGAATGCATCTCCAGCAAGAACACTTCAGCTGAGCATCT 70
Db 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGATCTCCAGCTGAGGACCTCCAGCGCGGATCTCCAGCCAGGATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyAlaArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCAGGATCTCCAGCTGAGGACCTCCAGCGCGGATCTCCAGCCAGGATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyAlaArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGTACACCTCCAGCGCGGATCTCCAGCGCGG-----GCATCT 235
Db 61 ProAlaGlyThrProProGlyAlaGlnAlaSerProGlyAlaArgAlaSerProAlaGlnAlaSer 80
QY 236 CCAGCCAGGATCTCCAGCGCGGATCTCCAGCTTGGCATCACTTCCAGGTCTCA 295
Db 81 ProAlaGlnAlaSerProAlaArgAlaSerProAlaGlnAlaSerLeuSerArgSerSer 100


```

QY 296 TCCGGAGTCATCATCCGAGGTACGCTCGGTGACAACTCCCAACGAGGTGAC 355
Db 101 SerOlyrGserSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
QY 356 CTGTGTAGAGCAACACGAGTGGGGCTGTATCCATCCGATCATCTCTCCAGGTGACGA 415
Db 121 LeuValAlrGAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
QY 416 CCGACACCGAGGGGACCAAGGAGAGAGCCGAGTACGAGCTCGCCAGTTACCTGGCGG 475
Db 141 ProAlaThrArgAlaThrArgGlySerProGlyThrSerLeuProLysPheThrTyrArg 160
QY 476 GAGGGGCAAGAGAGAGTACCGCTCATCGAGTGGGCTGCTCTCATTTAGCCCTGATGAT 535
Db 161 GluGlyGlnGlySerGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValVal 180
QY 536 TCGCTCATCATCTCTTCCAGTTCTGACAGGGCCACACAGGGATCAAGTACAGAGACAG 595
Db 181 SerLeuIleIleLeuPheGlnPheThrGlnGlyIleArgTyrLeuGlyGln 200
QY 596 AGGAGAGAGCTGTCCCAAGCAACGCTGTTCGCTGACCGGGGTGGAGTGCAGACTGAG 655
Db 201 ArgGlySerCysProLysHisAlaValAlaArgCysArgGlyValValAspCysLysLeuLys 220
QY 656 AGTACAGAGCTGGGCTCGTGAAGGTTTGACTGGAGACAAGTCTGTAAATCTACTCT 715
Db 221 SerThrGlyLeuGlyCysValAlaArgPheAspThrAspLysSerLeuLeuLysIleTyrSer 240
QY 716 GGGTCTCTCCATCATGAGTCTTCCATCTGTAGACAGAACTGGAATGACTCTACTCAGAG 775
Db 241 GlySerSerHisGlnTyrLeuProIleCysSerSerAsnTyrAsnAspSerTyrSerGln 260
QY 776 AAGACCTGGCCAGAGCTGGGTTTGAGAGTGTACACCGGACACCGAGGTTGGCCACAG 835
Db 261 LysThrCysGlnGlnLeuGlyPheGlySerAlaHisAspGlnThrGlnValAlaHisArg 280
QY 836 GATTTTGCACACAGCTTCTCAATCTTGAGATACAACTCCACATCCAGAGAAAGCTCCAC 895
Db 281 AspPheAlaAsnSerPheSerIleLeuArgGlyThrAsnSerThrIleGlnGlnSerLeuHis 300
QY 896 AGGTGTAAATGCCCTTCCACGGGTATATCTCCCTCAAGTGTCCCACTGGACGTAGAG 955
Db 301 ArgSerGlyCysProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
QY 956 GGCATGACCGGGGGAGTGTGGGAGGGGCGTGGCTCGGATAGCAAGTGGCTTGAGCA 1015
Db 321 AlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTyrProTyrGln 340
QY 1016 GTGAGTGTGCACTTCGGACACCACTCATGTGAGAGGACGCTCATTTGACGCCAGTGG 1075
Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTyr 360
QY 1076 GTGCTCACTGCGGCCACTGCTTCTTGTCGACCCGGAGAAAGTCTGGAGGCTGGAAG 1135
Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgLysValLeuGlnGlyTyrLys 380
QY 1136 GTGTAAGCGGGGACACAGCAACTGCAACAGTTGCTGAGGACGCTCATTTGSCGAGATC 1195
Db 381 ValTyrAlaGlyThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIle 400
QY 1196 ATCATCAACAGCAATTACACCGATAGAGAGAGCACTATGACATGCGCTCATGCGCTG 1255
Db 401 ThrIleAsnSerAsnTyrThrAspGlyGlnAspArgTyrAspIleAlaLeuMetArgLeu 420
QY 1256 TCCAAAGCCCTGACCCCTGTGCTGCTGACATCCACCTGCTTGCCTCCCATGATGAGACAG 1315
Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGln 440
QY 1316 ACCTTTAGCCTCATGAGACTGTGATGATCAAGGCTTTGGCAGAGACGAGGAGACAGAT 1375
Db 441 ThrPheSerLeuAsnGlnThrCysThrIleThrGlyPheGlyLysThrArgGlnThrAsp 460

```

```

QY 1376 GACAGACATCCCTTCTCCGGGAGTGCAGGTCAATCTCATGCACTTCAAGAAATGC 1435
Db 461 AspLysThrSerProPheLeuArgGlyValGlnValAsnLeuIleAspPheLysCys 480
QY 1436 AATGACTACTTGGTCTATGACAGTTTACCTTAACCCAAAGATGATGTGTGCTGGGGACCTT 1495
Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeu 500
QY 1496 CGTGGGGGACAGAGCTCTGCGCAGGGAGACAGCGGGGGGCTCTGTCTGTGACACAGAC 1555
Db 501 ArgGlyGlyAlaArgAspSerCysGlnGlyAspSerGlyLysProLeuValCysGlnGlnAsn 520
QY 1556 AACCGTGTATACCTGGCAGGTGTCCACCAAGCTGGGGCACAGGCTGTGGCCAGAGAAACAA 1615
Db 521 AsnArgThrTyrLeuAlaGlyValThrSerTyrGlyThrGlyCysGlyGlnArgAsnLys 540
QY 1616 CTTGGTGTATACCAAAAGTGACAGAAATTCTTCCCTGGATTTTACAGCAAGATGGAG 1672
Db 541 ProGlyValTyrThrLysValThrGluValLeuProTyrIleTyrSerLysMetGln 559

```

Search completed: September 16, 2006, 01:34:33
 Job time : 333 secs


```

OY 669 GACAAAGCTCTGCTTAATAATCTACTCTGGGCTCTCCCATCTAGTGGCTTCCCATCTGTAGC 748
Db 59 AlaaprlaahgyleMetValPheaprlaythrgluylYthrtprargyleuucysSer 78
OY 749 AGCAACTGGAATGACTCTCTACTCAGAGAAGACTCCGACAGCTGGATTTTCGAGAGTCT 808
Db 79 SeratgserAenllaargValaagIyLeuSerCySgIugIuMetGlyPheIeunargla 98
OY 809 CACCGAGAACCGAGAGTTGGCCCAAGGAT-----TTT 841
Db 99 LeuthrhiAseSgIuLeuapryalargThralagIyalaAenGlythSerGlyPhehe 118
OY 842 GCCAACAACCTTCTCAATCTTGAGATCAAACTCCACATCCAGAAAGACTCCACAGTCT 901
Db 119 CysValaaprgIugIyArgIeuProthiAthrGlnaargyleuGluValIleSerValCyS 138
OY 902 GAATGCCCTTCCGACCGGATATATCTCCCTCCAGTTCCTCCACTGGACTGAGAGGACTG 961
Db 139 AapCySproArgIyArgPheIeunlaaIaIeCySgInaapCySgIyArgargIyLeu 158
OY 962 ---ACGGGGCGGATGTGGAGAGGGGCGCTGGCTCGATAGCAAGTGGCTTGGGAAAGTG 1018
Db 159 ProValaapargIleValGlyyargAapThrSerIeunGlyArgtrProthrlaIVal 178
OY 1019 AGTGTGCACTTTCGGGACACCCACCAATCTGTGGAGGACGCTCATTTGAGCCGACAGGGTG 1078
Db 179 SerIeunargIyAaprgIyAlaIhIeIeucySgIyGlySerIeunSerGlyAapThrlaI 198
OY 1079 CTCACTGCCGCGCACTGCTTCTCTGTGACCCGGGAGAAAGTCTTGAGAGGCTGAGAGTG 1138
Db 199 LeuthrhlAaIaIhIeCyS---PheProGluArgaanaargValaIeSerArgThrgIyAl 217
OY 1139 TACCGGGGACC-----AGCAACTGCACCAAGTTGCTGTGAGGACAGCTCCATTGCC 1188
Db 218 PhealagIyAlaValaIaIaIaIeSerProthIeSgIyLeu-----GlnIeugIyAlaGln 235
OY 1190 GAGATCATCATCAACAGCAATTAC-----ACGAGTAGGAGGAGAGCAAC 1231
Db 236 AlaValaIaIyThrlaSgIyLytyrIeuProPheargAapProAenSerGlnIuIuAenSer 255
OY 1232 TATGACATCGGCGCTCATCGGCTGTGCAAGCCCTCGACCTGTGCGCTACATCCACCT 1291
Db 256 AenaprlleAlaIeIuValhIeIeunSerProIeuProIeuThrlGlytrIleGlnPro 275
OY 1292 GCTTGCTCTCCCATGATGAGACAGACCTTTAGCTCAATGAGACCTGCTGATCACAGCC 1351
Db 276 ValCySleuProAlaIaIaGlyGlnaIaIeIuValaaprgIyLySleCySthxValaThrlGly 295
OY 1352 TTTGGCAAGACCGAGGAGACAGATGACAAACATCCCTCCCTCCGGGAGGTGCAAGTTC 1411
Db 296 TrpGlyAenThrlGlytrIytrIeGlnaIaGly---ValleGlnIuIuAargVal 314
OY 1412 AATCTCATGCACTTCAAGAAATGCAATGACTACTGTGCTATGACAGTACTTACCTTCA 1471
Db 315 ProIleIleSerAenIaapryalCySleuGlnIyAlaapryeIyGlyAaenGlnIleuSerPro 334
OY 1472 AGATGATGTGTGCTGGAGACCTTGTGAGGAGGACAGACTCTGCTCAGAGGACAGCGAG 1531
Db 335 LysMetPheCySAlaGlytrProGlnIugIyIleAapAlaCyGlnIugIyAapSerGly 354
OY 1532 GGGCTCTTGTCTGTAG-----CAGAACACCGGTGTACTGTGCAAGTGTCTC 1579
Db 355 GlyProPheValCySgIuAapSerIleSerArgIythrProArgIytrArgIeucySgIyIle 374
OY 1580 ACCACTGTGGGACACAGGCTGTGGCCAGAGAAACAACCTGTGGGTGTACACCAAATGACA 1639
Db 375 ValSerTrpGlythrlCySAlaIeunlaGlnIuSerProGlyValIyThrluValSer 394
OY 1640 GAAGTTCTTCCCTGGATTATACAGCAAGATGAGAAC 1675
Db 395 AapPheargGluTrpIlePheGlnaIaIeIuSerThr 406

```

```

QY 1214 -----ACCGATGAGGAGGACGACTATGATGCGCTCCATGCGGCTTCCAAAGCCCTG 1267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 AspProThr11leAspGluAsnSerAsnAsp11leAlaValHisLeuSerSerLeu 266
QY 1268 ACCGTGTCGCGTCACATCCACCTCTTGCTGCCCTCCCATGATGACGACCTTTAGCTTC 1327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 ProLeuThrGluTrp11leGlnProValCysLeuProAla1aGlyGlnAlaLeuValAsp 286
QY 1328 AATGAGACCTGCTGGATCAGAGCTTTGGCAAGCCGAGACAGATGACAGACATCC 1387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 GlyValValCysThrValThrGlyTrpGlyAsnThr---GlnPheTrpGlyGlnAla 305
QY 1388 CCGTTCGCGGAGGAGGATGATCATGATGACCTTCAAGAAATGATGACTACTTGG 1447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 ValValLeuGlnGlnAlaAlaGValPro11leSerAsnGlnValCysAsnSerProAsp 325
QY 1448 GTCTATGACAGATTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGGGGGACGA 1507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 PheTrpGlyAsnGln11leLysProLysMetPheCysAlaGlyTrpProGluGlyGlyLe 345
QY 1508 GACTCTGCGCAAGGAGACAGCGGGGCGCTTGTCTGTGACAG-----AAC 1555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 AspAlaCysGlnGlnLysAspSerGlyGlyHisPheValCysGlnAspArg11leSerGlyThr 365
QY 1556 AACCGCTGTACCTCGGACAGGTGTACACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAA 1615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 SerArgTrpArgLeuCysGly11leValSerTrp11leThrGlyCysAlaLeuAlaArgLys 385
QY 1616 CTTGCTGTGATACCAAGTGAAGTCTTCTCCCTGATTTTACAGACAGATGAGAGAGC 1675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 ProGlyVal11leThrLysVal11leAspPheArgGluTrp11lePheGlnAla11leLysThr 405

```

RESULT 3
A56318
enteropeptidase (EC 3.4.21.9) precursor [validated] - human

N:Alternate names: enterokinase
C:Species: Homo sapiens (man)
C>Date: 19-May-1995 #sequence_revision 09-Aug-1996 #text_change 09-Jul-2004
C/Accession: A56318; B43090
R:Kitamoto, Y.; Veille, R.A.; Donis-Keller, H.; Sadler, J.E.
Biochemistry 34, 4562-4568, 1995
A>Title: cDNA sequence and chromosomal localization of human enterokinase, the proteolytic
A/Reference number: A56318; MUID:95324679; PMID:7718557
A/Accession: A56318
A:Molecule type: mRNA
A/Residues: 1-1019 <KIT>
A/Cross-references: UNIPROT:P98073; UNIPARC:UPI000003PE65; GB:U09860; NID:G746412; PIDN:
R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A>Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease compo
A/Reference number: A43090; MUID:94329561; PMID:8052624
A/Accession: B43090
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A/Residues: 749-1019 <KIT>
A/Cross-references: UNIPARC:UPI0000172B0B; GB:U09860
C/Comment: The mechanism of association with the membrane of the intestinal brush border
ocated below) or with amino-terminal myristoylation of the heavy chain.
C/Genetics:
A/Gene: GDB:PRSS7
A/Cross-references: GDB:384083; OMIM:226200
C/Map position: 21q21-21q21
C/Comment: Mature enteropeptidase is variously reported to contain two (heavy and light)
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involy
ducts.
C/Function:
A/Description: cleaves activation peptide from trypsinogen to produce active trypsin
A/Pathway: intestinal digestive hydrolyase cascade
C/Superfamily: enteropeptidase; Ctr/Cis repeat homology; LDL receptor ligand-binding rep
C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
F:1-784/Product: enteropeptidase heavy chain #status predicted <HCH>
F:22-38/Domain: transmembrane #status predicted <TM>

F:184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:342-504/Domain: MM homology <MM>
F:526-631/Domain: Ctr/Cis repeat homology <CTR>
F:643-677/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:678-783/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRCT
F:785-1019/Product: enteropeptidase light chain #status predicted <LCH>
F:785-1014/Domain: trypsin homology <TRY>
F:116-147, 179, 328, 335, 388, 440, 470, 503, 534, 630, 682, 706, 725, 848, 887, 909, 949/Binding site: <
F:772-896, 810-825, 910-977, 941-956, 967-995/Disulfide bonds: #status predicted
F:825, 876, 971/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 1,01e-24 Length: 1019
Score: 580.00 Matches: 141
Percent Similarity: 47.5% Conservative: 56
Best Local Similarity: 34.0% Mismatches: 159
Query Match: 17.4% Indels: 59
DB: Gaps: 14

US-10-806-370-11 (1-1748) x A56318 (1-1019)

```

QY 529 GGTGGTTTC-----GCTCATCAT-----CTTTCAGTTCTG 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 GlyGlyPheLysAlaAsnPheThrThrGlyTrpHisLeuGlyLeProGluProCysLys 644
QY 562 GCAGGGCCACACAGGATCATGATCAAGAGACAGAGGAGGATGTCCTCCAAAGACGCTGT 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 645 Ala-----AspHisPheGlnCysLysAsnGlyGln---CysValProLeuValAs 660
QY 622 TCCCTGTGACGAGGATGATGACTCAAGCTGGAAGATGACGAGCTGGCTGCTGAGATT 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 nLeuCysAspGlyHisLeuHisCysGlnAspGlySerAspGlyAlaAspCysValAlaArg 680
QY 682 TGACTGGAGCAAGCTCTGTTAAATCTACTGTGGCTCCCATCAG----- 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 680 e-----PheAsnGlyThrThrAsnAsnAsnGlyLeuVal 691
QY 731 -----TGCTTCCATCTGTAGACAGACATGTGAAATGACTCTTA 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 ArgPheArg11leGlnSer11leTrpHisThrAlaCysAlaGlnAsnThrThrGln11 711
QY 769 CTCAGAGAAAGACCTGCGCAGCAGCTGGCTTGCAGAGTGTCTACCGGACCAACCGAGTTGC 828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 711 eSerAsnAspValCysGlnLeuLeuGlySerGlyAsnSerSerLysPro11lePhe 731
QY 829 CCACAGGATTTTGCACACAGCTTCTCATCTTGATACAACTCCACCATCCAGAGAAAG 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 eSerThrAsp-----GlyGlyProPheValLysLeuAsnThrAlaProAspGlyHis 748
QY 889 CCTTC-----CACAGCTGAAATGCTTCCACCGGTATATCTCCCTCCAGTTTC 939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 748 sLeu11leLeuThrProSerGlnGlnCysLeuGlnAspSerLeu11leArgLeuGlnCysAs 768
QY 940 CCAC-----TGCGGACTGAGGGCC-----ATGACCGGGGAGATCGGGAGG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 768 nHisLysSerCysGlyLysLeuLysLeuAlaAlaGlnAsp11leThrProLys11leValGly 788
QY 982 GCGCTGCGCTCGAGATGACAGTGGCTTGGCAAGTGAATGCTGCACTTGGCACCAACCA 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 ySerAsnAlaLysGlnGlyLysAlaTrpProTrpValValGlyLeuTrpTrpGlyArgLe 808
QY 1042 CATCTGTGAGGACGCTCATTTGACGCCACGCTGGGCTCTACTGCGCCGCACTGCTTCTT 1101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 uLeuCysGlyAlaSerLeuValSerSerAspTrpLeuValSerAla1aHisCysVal11 828
QY 1102 CGTGACCGCGGAGAGAGTCTTGAGGGCTGGAAGGTGATGACGGGCGCACCAACTGCA 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 828 r---GlyArgAsnLeuGlnProSerLysTrpThrAla11leLeuGly-----LeuHis 844
QY 1162 CCAATTGCTGAGACAGCTTC-----ATTGCGAGATCATCAT 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 sMetLysSerAsnLeuThrSerProGlnThrValProArgLeu11leAsp11leVal11 864

```

[illegible]

Db 436 Phe-----AspGlyLeuProLeuGlnAspValTrpArgIle 447
 QY 1139 TAGCGGGGACACCACTTCACCACTTGCTGCTAG-----GCAGCTCCATTGCC 1189
 Db 448 TyrSerGlyIleLeuAsnLeuSerAspIleThrIysAspThrProPheSerGlnIleIlys 467
 QY 1190 GAGATCATCATCAGACCAATTACACCGATGAGGAGGACGATGACATGACGCTCATG 1249
 Db 468 GlnIleIleIleIleIleGlnAsnTrpIleValSerGlnGlnAsnIleAspIleIleLeuIle 487
 QY 1250 CGAGTGTCCAGACCCCTGACCCCTGTCGCTCATATCCACCTGCTGCTCCCATGAT 1309
 Db 488 LysLeuGlnIleProLeuAsnTrpThrGlnPheGlnLysProIleCysLeuProSerIlys 507
 QY 1310 GGAACAGACCTTACCTTCATGAGACCTGCTGATCAGAGCTTGGCAACACAGGAG 1369
 Db 508 GlyAspThrSerThrIleTyrThrAsnIleTrpValThrGlyTrpGlyPheSerIleGln 527
 QY 1370 ACAGATGACAAAGACATCCCTCTCCGAGGAGGAGGACGATCATGCTGATGATCAG 1429
 Db 528 LysGlyGlu---IleGlnAsnIleLeuGlnIleValAsnIleProLeuValThrAsnGln 546
 QY 1430 AATGCAATGACTACTGCTGATGACAGATTAC---CTTACCCCAAGAGATGATGCTGCT 1486
 Db 547 GluCysGlnIleValArg-----TyrGlnAspTrpIleThrGlnArgMetValCysAla 564
 QY 1487 GGGACCTTCTGGGGGAGAGACTCTGCGACGAGAGACAGCGGGGCTCTTGTCTGT 1546
 Db 565 GlyTyrIleGlnGlyIleGlyIleAspAlaCysIleGlyAspSerGlyIleProLeuValCys 584
 QY 1547 GAGCAAGAACACCCCTGCTACTGCGACGCTGACGCTGCGACGCTGCGACGCTGCGACG 1606
 Db 585 LysHisAsnGlyMetTrpArgLeuValGlyIleThrSerTrpGlyGlyCysAlaArg 604
 QY 1607 AGAACAACAACTGGTGTGTATACCAAGATGACAGAGTTCTCCCTGATTTACAGCAAG 1666
 Db 605 ArgGlnGlnProGlyValTyrThrIleValAlaGlnIleTrpMetAspTrpIleLeuGlnIlys 624
 QY 1667 ATGGAGAGC 1675
 Db 625 ThrGlnSer 627
 RESULT 7
 KORTPL
 plasma kallikrein (EC 3.4.21.34) precursor - rat
 N/Alternate names: Fletcher factor; kininogenin; serum kallikrein
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text_change 05-Oct-2004
 C/Accession: A39180; A33320; S06851; I53041; S06852
 R/Beaubien, G.; Kosinski-Chupin, I.; Matter, M.G.; Mbikay, M.; Chretien, M.; Seidah, N.G.
 A/Title: Gene structure and chromosomal localization of plasma kallikrein.
 A/Reference number: A39180; MUID:91129236; PMID:1993180
 A/Accession: A39180
 A/Molecule type: DNA
 A/Residues: 1-638 <BNA>
 A/Cross-references: UNIPROT:P14272; UNIPARC:UPI000004ECC9; GB:J05315
 A/Note: the authors translated the codon GAG for residue 81 as Gln
 R/Seidah, N.G.; Lidenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazure
 DNA 8, 563-574, 1989
 A/Title: The cDNA structure of rat plasma kallikrein.
 A/Reference number: A33320; MUID:90091743; PMID:2598771
 A/Accession: A33320
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-638 <SEI>
 A/Cross-references: UNIPARC:UPI000004ECC9; GB:M30282; NID:G205010; PIDN:AAA4163.1; PID:
 A/Note: part of this sequence, including the amino ends of both the heavy and light chain
 R/Paguin, J.; Benjannet, S.; Sawyer, N.; Lazure, C.; Chretien, M.; Seidah, N.G.
 Biochim. Biophys. Acta 999, 103-110, 1989
 A/Title: Rat plasma kallikrein: purification, NH(2)-terminal sequencing and development
 A/Reference number: S06851; MUID:90089457; PMID:2597701
 A/Accession: S06851

A/Molecule type: protein
 A/Residues: 20-45;391-413 <PAO>
 A/Cross-references: UNIPARC:UPI0000172B57; UNIPARC:UPI0000172B58
 R/Seidah, N.G.; Lidenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazure
 DNA Cell Biol. 8, 563-574, 1989
 A/Title: The cDNA structure of rat plasma kallikrein.
 A/Reference number: I53041
 A/Accession: I53041
 A/Status: translated from GB/EMBL/DBD
 A/Molecule type: mRNA
 A/Residues: 1-638 <RES>
 A/Cross-references: UNIPARC:UPI000004ECC9; GB:M58590; NID:G206721; PIDN:AAA42069.1; PID:
 C/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w/
 C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 115
 are linked by one or more disulfide bonds.
 C/Genetics:
 A/Gene: PK
 C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
 C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; inflan
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-390/Product: plasma kallikrein heavy chain #status experimental <MAT1>
 F:20-109/Domain: apple repeat <AP1>
 F:110-199/Domain: apple repeat <AP2>
 F:200-289/Domain: apple repeat <AP3>
 F:291-380/Domain: apple repeat <AP4>
 F:391-638/Product: plasma kallikrein light chain #status experimental <MAT2>
 F:391-621/Domain: trypsin homology <TRY>
 F:121-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322-
 F:127,215,308,453,459,494/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:396/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:433,483,578/Active site: His, Asp, Ser #status predicted
 Alignment Scores:
 Pred. No.: 4,7e-24 Length: 638
 Score: 568.00 Matches: 144
 Percent Similarity: 44.8% Conservative: 62
 Best local Similarity: 31.3% Mismatches: 131
 Query Match: 17.0% Indels: 123
 DB: 1 Gaps: 18
 US-10-806-370-11 (1-1748) x KORTPL (1-638)
 QY 403 TGCCAGTCAACAGACCAACAGGACAGGAGGAGCCAGGACGAGTACGAGCTGCCCA 462
 Db 256 CyheheleuylshrsrlyserglyArgProSerProPro-----IlelleGln 272
 QY 463 GTTCACCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 522
 Db 273 GlnAsnIleValSerGlyTyrSerLeuPheThrCysArgIleAlaArgProGluPro-Cy 292
 QY 523 TGCCCTGGTGTTCGCTCATCATCTCTTCGAG-----TTCTG 561
 Db 292 ShlsPheIleIleTyrSerGlyValAlaPheGlnGlyGlnIleuAsnAlaThrPheVa 312
 QY 562 GCAGGCGCACAGAGGATGACGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
 Db 312 IghlyAlaIleAspAlaCys-----GlnGlnThrCysThrIys---ThrI 326
 QY 622 TCGCTGTAGCGGGGTGTG-----GACTGACAGCTGAGAGTGAAGTGAAGA 663
 Db 326 eArgCysGlnPhePheThrTyrSerLeuLeuProGlnAspIleValIleGlu----- 343
 QY 664 GCTGGGCTGC-----GTGAGTTTGACTGGAGCAAGCTCTGCTTAATC----- 709
 Db 344 ---GlyCysIleCysSerLeuArgLeuSerThrAspGlySerProThrArgIleThrTy 362
 QY 710 ---TACTGGGGGCTCCCATCATGAGGCTCCCATGATGAGCAAGTGAATGACTC 765
 Db 362 rgAlaIleGlnIleSer-----G1 369
 QY 766 CTACTGAGAGAGACCTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 825
 Db 369 yTyrSerLeuArgIleuValIle-----GluSerSerAspCysThrThrIysI 387


```

QY      826 TGCCACAGGAGATTGTCACACAGCTTCTCAATCTGAGATACAACTCCACCATCCAGGA 885
D8      387 e-----
QY      886 AAGCTCCACAGGCTGTAATGCCCTTCCACGGATATCTCCCTCCAGTGTCCCACTG 945
D8      387 -----
QY      946 CGAGTCAGGCGCATGACCGGCGGATGTCGAGAGGCGCTGCGCTGCATGACCAAGT 1005
D8      388 -----
QY      1006 GCCTTGGCAAGTAGTGTGCAC-----TTCGGCACCAACCAATCTGTGAGGCGAC 1056
D8      402 pPrtGlnValSerLeuGlnValLysLeuValSerGlnAsnHisMetCysGlyGlySer 422
QY      1057 GCTCATTAAGCCCGAGGCGGCTGTCACGCGGCCCATGCTTCTTCGAGACCGCGGAGAA 1116
D8      422 rTleIleGlyValGlnTrpIleLeuThrAlaIleHisCysPhe----- 436
QY      1117 GATCTCGAGAGGC-----TGAAGAGTGAACGCGGCGACACCAACT 1158
D8      437 -----
QY      1159 GCACCAAGTTCCTGAGGCA-----GCTTCATTCGCGAGATCATCATCAACAGCAA 1209
D8      454 uSerGluIleThrAsnLysThrProPheSerSerIleGlyGluLeuIleHisGlnLys 474
QY      1210 TTACACGATGAGAGAGACGATGACATGCGCTCATGCGGCTGTCACCGCCCTGAC 1269
D8      474 rTyrLysMetSerGlnGlySerTyrAspIleAlaLeuIleLysLeuGlnThrProLeuIle 494
QY      1270 CTTGTCGCTCATGATCCACCTGCTGCTCCCATGATGATGAGACGACTTAACTCAA 1329
D8      494 nTyrThrGluPheGlnLysPheIleCysLeuProSerIleLysAlaSerHisAsnThrIleTyr 514
QY      1330 TGAGACCTGCTGATACACAGGCTTTGGCAGACGAGGAGACAGATGACATCCCC 1389
D8      514 rThrAsnCysSerValThrGlyTyrGlyTyrThrLysGlnLysArgLysGln-----ThrGlnAs 533
QY      1390 CTTCTCCGAGGAGTGACGAGCTCATTCATGATGATGATGATGATGATGATGATGATGAT 1447
D8      533 nIleLeuGlnLysAlaThrIleProLeuValProAsnGlnGlnLysGlnLysLysTyrAr 553
QY      1438 TGACTACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
D8      553 gAspTyrValIle-----ThrLysGlnMetIleCysAlaGlyTyrLysGln 568
QY      1498 TGGGGGACAGAGATCTGCGACAGGAGACAGCGGGGGGCTTGTCTGTGACAGAAACA 1557
D8      568 uGlyGlyIleAspAlaCysLysGlyAspSerGlyGlyProLeuValCysLysHisSerGln 588
QY      1558 CCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1617
D8      588 yAspTyrGlnLeuValGlyIleHisSerTyrGlyGlnGlyCysAlaArgLysGlnGlnPr 608
QY      1618 TGGGTGTATACCAAGATGACAGAGAGTCTTCCCTGATTTACAGCAAGATGAGAGAC 1675
D8      608 oGlyValTyrThrLysValAlaGlyTyrIleAspTyrIleLeuGlnLysIleGlnSer 627

RESULT 8
A53663
Enteropeptidase (EC 3.4.21.9) precursor [validated] - pig
N/Alternate names: enterokinase
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 07-Oct-1994 #sequence_revision 09-Aug-1996 #text_change 28-Apr-2003
C/Accession: A53663
R/Matsuhashima, M.; Ichinose, M.; Yahagi, N.; Kakei, N.; Tsukada, S.; Miki, K.; Kurokawa,
J. Biol. Chem. 269, 19976-19982, 1994
A/Title: Structural characterization of porcine enteropeptidase.
A/Reference number: A53663; MUID:94327548; PMID:8051081
A/Accession: A53663

```

```

A/Molecule type: mRNA
A/Residues: 1-1034 <MAT>
A/Cross-references: UNIPARC:UPI0000172B0C; GB:D30799; NID:G505122; PIDN:BA06459.1; PID:6
A/Note: parts of this sequence, including the amino ends of three chains isolated from t
C/Comment: The mechanism of association with the membrane of the intestinal brush border
ated below) or with amino-terminal myristoylation of the heavy chain.
C/Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involve
C/Function:
A/Description: cleaves activation peptide from trypsinogen to produce active trypsin
A/Pathway: intestinal digestive hydrolyase cascade
C/Superfamily: enteropeptidase; Ctr/Cts repeat homology; LDL receptor ligand-binding repe
C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
F/22-38/Domain: transmembrane #status predicted <TM>
F/52-117/Product: enteropeptidase mini chain #status predicted <MCH>
F/118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
F/119-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F/357-519/Domain: MAM homology <MAM>
F/541-646/Domain: Ctr/Cts repeat homology <CTR>
F/658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F/693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRCT
F/800-1034/Product: enteropeptidase light chain #status predicted <LCH>
F/800-1029/Domain: trypsin homology <TRY>
F/116,147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,902,966
F/787-911,925-841,925-992,956-971,982-1010/Disulfide bonds: #status predicted
F/840,891,986/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 2,13e-23 Length: 1034
Score: 556.00 Matches: 143
Percent Similarity: 46.7% Conservative: 56
Best Local Similarity: 33.6% Mismatches: 164
Query Match: 16.7% Indels: 63
DB: Gaps: 15

US-10-806-370-11 (1-1748) x A53663 (1-1034)
QY      475 GAGAGGCGCAGAGCA-----GTAACGCTCATCGGTCGCTGCTCTCTCATTCGCT 528
D8      640 GlyGlyPheLysValAlaAsnPheThrThrGlyTyrHisLeuGlyIleProGluProCys--- 658
QY      529 GGTGGTTGCGTATCATTCCTCTCCAGTTCTGCGACGCGCACAGGACATGACTTCAA 588
D8      659 -----LysGluAspAsnPheGln 664
QY      589 GAGACAGAGGAGAGAGTGTCCCAAGACAGCTGTTGCTGACCGGGGTGATGACGCA 648
D8      665 CysGlnAsnGlyGln-CysValLeuLeuValAsnLeuCysAspGlyPheSerHisCysGly 684
QY      649 GCTGAGAGTGACAGAGCTGGGCTGCGTGAAGTTTGAATGGGACAGATGCTGTTAAAT 708
D8      684 AspGlySerAspIleAlaHisCysValArgPhe----- 695
QY      709 CTACTCGGAGTCCTCCCATGAG-----TGGCT 735
D8      696 -LeuAsnGlyThrAlaAsnSerGlyLeuValGlnPheArgIleGlnSerIleTrpHis 715
QY      736 TCCCATGTGAGACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
D8      715 rThrAlaCysAlaGlnAsnTrpThrGlnThrSerAspAspAlaCysGlnLeuLeuGln 735
QY      796 TTTCGAGAGTGTCCACCGACCAACCGAGTTGCCCAAGGATTTGCCAAGCTTCTC 855
D8      735 yLeuGlyThrGlyAsn-----SerSerMetProPhePheSerSerGlyGln 750
QY      856 A-----ATCTGAGATATCACTCCACATCCAGGAAGCTC-----CACAGTC 900
D8      750 yGlyProPheValLysLeuAsnThrAlaProAsnGlySerLeuIleLeuThrAlaSerGln 770
QY      901 TGAATGCCCTTCCACGAGGATATATCTCCCTCCAGTGTCCAC-----TGCGGA----- 949
D8      770 uGlnCysPheGlnAspSerLeuIleLeuLeuGlnCysAsnHisLysSerCysGlyLysGly 790

```



```

Db      1002 -----GlyAlaHisMetProPheLeuGlnGluValArgIleIle 1017
Qy      1421 GACTTCAAGAAATGCAATGACTTACTGTCTATGAC--AGTTACTTATCCCAAGATG 1477
      1018 ProLeuGlnGlnCysGlnSerTyr-----PheAspMetLysThrIleThrAsnArgMet 1035
Qy      1478 ATGTGTCTGGGGGACTTTCGTCGGGCGAGAGACTTCTCCAGGAGACAGAGGGGGGCTT 1537
      1036 IleCysAlaGlyTyrGlnSerGlyThrValAspSerCysMetGlyAspSerGlyGlyPro 1055
Qy      1538 CTTCTCTGTGAGCAG--AACACCGCTGTACCTGGCAGGTGCACGAGTGGGGACA 1594
      1056 LeuValCysGlnGluAlaGProGlyGlnThrPheLeuPheGlyLeuThrSerTrpGlySer 1075
Qy      1595 GCGTGTGGCCAGAGA--AACAACTGTGTGTATACCAAGATGACAGAAATTTCTCC 1651
      1076 ValCysPheSerLysValLeuGlyProGlyValTyrSerAsnValSerTyrPheValGly 1095
Qy      1652 TGGATT 1657
      1096 TrpIle 1097

RESULT 11
JC5759
brain-specific serine proteinase (EC 3.4.21.-) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 31-Dec-2004
C/Accession: J05759
R/Yamamura, Y.; Yamashiro, K.; Tsunoka, N.; Nakazato, H.; Tsujimura, A.; Yamaguchi, N.
Biochem. Biophys. Res. Commun. 239, 386-392, 1997
A/Title: Molecular cloning of a novel brain-specific serine protease with a kringle-like
A/Reference number: J05759; MUID:98008848; PMID:9344839
A/Accession: J05759
A/Molecule type: mRNA
A/Residues: 1-761 <YAM>
A/Cross-references: UNIPARC:UPI0000175C77; DDBJ:D89871
A/Experimental source: brain
C/Superfamily: plasma hyaluronan-binding protein; scavenger receptor cysteine-rich domain
C/Keywords: glycoprotein; hydrolase; serine proteinase
F/85-157/Domain: kringle-like #status predicted <KR>
F/163-266/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F/166-266,273-372,386-486/Domain: scavenger receptor cysteine-rich #status predicted <SR>
F/513-516/Domain: furin binding #status predicted <TRY>
F/517-755/Domain: tryptsin homology <TRY>
F/93,521,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/562,612,711/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 1.64e-22 Length: 761
Score: 540.00 Matches: 194
Percent Similarity: 34.8% Conservative: 73
Best Local Similarity: 25.3% Mismatches: 240
Query Match: 16.2% Indels: 262
DB: 2 Gaps: 30

US-10-806-370-11 (1-1748) x J05759 (1-761)
Qy      59 GCTGAGCATCTCCAGCCAGCATCTCCAGTGGGACACT-----CCAGCCGG 109
      19 AlaArgAlaSerProValSerArgSerProLeuHisArgProHisProSerProArg 38
Qy      110 GCATTCAGCCAGCATCTCCAGCCAGCATCTCCAGTGGGACACT-----160
      39 SerGlnHisAlaHisTyrLeuProSerSerArgArgProArgTrpArgPhePro 58
Qy      161 -----CCGGCCGGGCACTCCAGCCAGCATCTCCAGTGGGACACT 205
      59 LeuProLeuArgIleProAlaAlaGlnArgProGlnValLeuSerThrGlyHisThrPro 78
Qy      206 -----CCAGCCGGGCACTCTCCAGC-----226
      79 ProThrIleProArgArgCysGlyAlaGlyGlnSerTrpGlyAsnAlaThrAsnLeuGly 98

```

```

Qy      227 -----CGGCACTCCAGCCAG 244
      99 ValProCysLeuHisTyrPheArgIleValProProPheLeuGlnLysSerProProAlaSer 118
Qy      245 GCATCTCCAGCCGGGCACTCCGGCTCCGATCACTTCCAGTTCCTGATCCGGAG 304
      119 TrpAlaGlnLeuLysArgIleGlnProHisAsnPheCysArgSerProArgGlySerGlyArg 138
Qy      304 -----304
      139 ProTrpCysPheTyrArgAsnAlaGlnGlyValAspTrpGlyTyrCysAspCysGly 158
      305 -----TCATCTCCGGCAGTCCAGCTCCG 328
Qy      159 GlnGlyProAlaLeuProValIleArgLeuValGlyLysAspSerGlyHisGlnGly---177
      329 GTGACAACTCCCAACAGAGTGTACTTGTAGACAAACAGAGTGGGGCTGTACC 388
      178 -----ArgValGlnLeuTyrHisAla-GlyGlnTrpGlyThrIleCys 191
Qy      389 ATCCGATCAT-----398
      191 AspAspGlnTrpAspAsnAlaAspAlaAspValIleCysArgGlnLeuGlyLeuSerGly 211
Qy      399 -CTCCTCCAGCTCAGCAGCAGCA-----CCAGGCCACAGGGAGAGC 442
      211 YIleAlaLysAlaTrpHisGlnAlaHisPheGlyGlnGlySerIleProIleLeuLeuAs 231
Qy      443 CCAGGTACGAGCCTG-----CCAAAGTTCACCTGGCG 474
      231 PGIValAlaArgCysThrGlyValGlnGlnMetSerIleGlnGlnCysProLysSerSerTrpGly 251
Qy      475 GGAG-----GGCAGAGCAGCTACCGCTCAGGCTGGCTCTCTCATTTGC 525
      251 YGlnHisAsnCysGlyHisLysGlnAspAlaGlyValSerCysValProLeuThrAspGly 271
Qy      526 CTTGTGTGTTTCCGTATCATCTCTCCAGTTCGGCAGGCGCACAGGATCAGGTA 585
      271 Y---ValIleArgLeuLysAlaGlyLysSerThrHisGlnLysArgLeuGlnValTyrTyr 290
Qy      586 CAAGGACAGAGGAGAGC--TGTCACAGCAGCTGTTGCTGTGACGGGGTGTGA 642
      290 LysValGlnTrpGlyThrValCysAspAspGlyTyrThrGlnMetAsnThrTyrValAl 310
Qy      643 CTGCAAGCTGAAGAGTACAGCTGGC-----670
      310 AcValArgLeuLeuGlyPheLysTyrGlyLysGlnSerSerValAsnHisPheAspGlySe 330
Qy      670 -----670
      330 AsnArgProIleTrpLeuAspAspValSerCysSerGlyLysGlnValSerPheIleGly 350
Qy      671 -TGGGTAGGTTTACTCG--GACAAGTCTTG--700
      350 nCysSerArgArgGlnTrpGlyArgHisAspCysSerHisArgGlnAspValGlyLeuTh 370
Qy      701 -----CTTAATCTACTC 714
      370 rCysTyrProAspSerAspGlyHisArgLeuSerProGlyPheProIleArgLeuValAs 390
Qy      715 TGGGTCTCCCAT-----CAGTGGCTTCCAT 741
      390 PGIValGlnAsnLysLysGlnGlyArgValGlnValPheValAsnGlyGlnTrpGlyThrI 410
Qy      742 CTGTAGAGCAACTGGAATGACTCTACTCAGAGAAACCTGCAGAGAGTGGTTTGA 801
      410 eCysAspAspGlyTyrThrAspLysHisAlaAlaValIleCysArgGlnLeuGlyTyrLys 430
Qy      802 GAGTGTACCCGCAACAGAGTTCGCCACAGGAGATTTGCCAAACGCTTCAATCTT 861
      430 sGlyProAlaArgAlaArgThrMetAlaTyr-----PheGlyGlnGlyLysGlyProI 448

```

```

QY      862 GAGATACACTCCAC-----ATCCAGCA 885
      :      :
Db     448 AHEMetAsperNallYsCystrNglyAaanglyAlaLeuAlaApcyValYsgl 468
QY      886 AAGCTCCACAGGCTGTAATGCCCTCCACGGGTATATCTCCCTCCAGT----- 937
      :      :
Db     468 nAspIleGlyAArgHIsaancysAArgHIsaSerGluAspAlaGlyAlIleCyAspTyrLe 488
QY      938 -----TCCCATGCGGACGAG 954
      :      :
Db     488 uGluLylAlaAsaSerSerSerGlyAsnLysGluMeleuSerSerGlyCySglYleuAr 508
QY      955 GGCCATG-----ACCGGGCGGATCGTGGAGGGCGGCTGCGCTCGATGACAAAGT 1005
      :      :
Db     508 gLeuLeuHIsaArgArgInLysArgIleIleGlyGlyAasnSerLeuAArgIAlaTr 528
QY      1006 GCCTTGGCAAGTGAAGTCTGCATTCGGACCAACCCAC-----ATCTGGG 1050
      :      :
Db     528 pProTyrpIlnAlaSerLeuAArgLysArgSerAlaHIsGlyAAspGlyArgLeuLysG 548
QY      1051 AGGCAGCTCATTTGACGCCGACGCTGCTACATGCGCCGCTGCTTCTTCTGAGACCG 1110
      :      :
Db     548 yAlaThrLeuLeuSerSerCystrpValLeuThrAlaAlaHIsCySpc----- 564
QY      1111 GGAGAAAGTCTGAGAGGCTGAAAGGTGTACCGCGGACCAAGCAACTGCAC----- 1162
      :      :
Db     565 -----LysArgTyrGlyAasnAsnSerAArgSerTyrAlaValAr 577
QY      1163 -----CAGTTCCTTGAAGCAGCTTCATTCGCGA 1191
      :      :
Db     577 gValGlyAAspTyrHIsThrLeuValProGluGluPheGluInGluIleGlyAlaGlnG 597
QY      1192 GATCATCTCAACAGCAATTAACACGATGAGAGAGAGACGATGACATGCGCTTCATCG 1251
      :      :
Db     597 nIleValIleHIsaArgAsnTyrAArgProAspAArgSerAAspTyrAspIleAlaLeuValAr 617
QY      1252 GCTGTCCAAAGCC-----CTGACCTGTGCGCTGCATCCACCTGCTTGGCT 1299
      :      :
Db     617 gLeuGlnGlyProGluGluInGlyCysAlaAArgLeuSerThrHIsValLeuProAlaCyLe 637
QY      1300 CCCCATG-----CATGACAGACCTTGAACCTCAATGACACCTGCTGATACAGG 1350
      :      :
Db     637 uProLeuTyrAArgGluAArgProGluLysThrAlaSerAsn-----CysHIsIleThrG 655
QY      1351 CTTTGGACAGACGAGGAGACAGATGACAGACATCCCTTCCCTCCGGAGGTGACAGT 1410
      :      :
Db     655 yTyrGlyAAspHrNglyAArgAlaTyrSerAArgThr-----LeuGlnGlnAlaAlaVa 672
QY      1411 CAATCTCATGACTTCAAGAAATGCAATGACTACTTGTCTTATGACAGTTACCTTACCC 1470
      :      :
Db     672 lProLeuLeuProLysAArgPheCyLysGluArg-----TyrLysGlyLeuPheThrG 690
QY      1471 AAGATGATGTGTCTGAGGACCTTCGTTGGGGGCA-----GACCTTCCTCCAGGAGA 1524
      :      :
Db     690 yAArgMetLeuCySAlaGlyAasnLeuGlnGluAAspAAsnAArgValAAspSerCySgInGlyAs 710
QY      1525 CAGGAGGGGGCTTGTCTGTGAGCAG-----AACACCGCTGTACTGACGAGGAGTGTAC 1581
      :      :
Db     710 pSerGlyGlyProLeuMetCySgInLysProAspGluSerTrpValValTyrGlyValTh 730
QY      1582 CAGCTGGGGACAGAGCTGTGGCCAGAGAAACAACTGTGTGTATACCAAAAGTGACAGA 1641
      :      :
Db     730 rSerTrpGlyTyrGlyCySgIlyValAAspThrProGlyValTyrThrAArgValProAl 750
QY      1642 AGTTCTTCCCTGAGATTACAGC 1663
      :      :
Db     750 aPheValProTrpIleLysSer 757

```

```

RESULT 12
A47547
serine proteinase stubble-stubloid (BC 3.4.21.-) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

```

```

C:Accession: A47547
R:Appl. L.F.; Prot. M.; Abu-Shumays, R.; Hammond, A.; Garbe, J.C.; Friele, D.; Friele
Proc. Natl. Acad. Sci. U.S.A. 90, 4937-4941, 1993
A:Title: The Drosophila stubble-stubloid gene encodes an apparent transmembrane serine I
A:Reference number: A47547; MUID:93281671; EMBL:7685111
A:Accession: A47547
A:Structure: preliminary
A:Molecule type: mRNA
A:Residues: 1-786 <APP>
A:Cross-references: UNIPROT:Q05319; UNIPARC:UP1000013615A; GB:L11451; MID:G158511; PIDN:
C:Genetic:
A:Gene: Sb-bd
A:Cross-references: FlyBase:FBgn000319
C:Superfamily: serine proteinase stubble-stubloid; trypsin homology
C:Keywords: hydrolase; serine proteinase; transmembrane protein
F:61-77/Domain: transmembrane #status predicted <TWN>
F:543-781/Domain: trypsin homology <TRY>

Alignment Scores:
Pred. No.: 1,74e-22 Length: 786
Score: 539.50 Matches: 181
Percent Similarity: 46.7% Conservative: 92
Best Local Similarity: 31.0% Mismatches: 224
Query Match: 16.2% Indels: 88
DB: Gaps: 23

US-10-806-370-11 (1-1748) x A47547 (1-786)

QY      8 ACCATGAGAGGACACCCAGGGAATGATCTCCAGCAAGAACACTTGACGTGAGCA 67
      :      :
Db     254 ThrThrAlaSerAAspLeuHIsGlySerAlaSerHIsProSerSerSerSerSer 273
QY      68 TCTTCACCCAGGAGATCTCCAGCTGAGAGACCTTCAGCGGAGCATCTCCAGCAGCA 127
      :      :
Db     274 SerSerSerSerProAan-serIleTrpHIsThrSerThrGlnGlnGlnGlnH 293
QY      128 TCTTCACCCAGGAGCATCTCCAGCTGAGC---ACCTCCGGGCGGAGC---ATCTCCAGCC 181
      :      :
Db     293 sGlnGlnAlaGlnGlnAlaAsnHIsTrpGlnMetThrTrpGluProSerPheIleThrLys 313
QY      182 CAGGCACTCTCCAGTGTATACCTTCAGCGCGGAGCATCT---CGAGCGCGGAGCATCTCCA 238
      :      :
Db     313 roAArgProThrGlyTyr-ThrLysProGlyIleValAAsnLeuProMetProAlaArgPro 332
QY      239 GCCCAGGACCTCCAGCGCGGAGCATCTCCGAGCTGAGCATCATCTTCCAGGTCTCATCC 298
      :      :
Db     333 SerLysProSerLysProThrLysLysProIleVal-----TyrAAspSerProPro 350
QY      299 GGCAGTCAATCATCCGACAGGTCAAGCTCGGTGACAACTCCCAACCAAGGTACCTT 358
      :      :
Db     351 ProProSerSerValProProSerThrSerThrSerThrSerThrSerThrSerLeuIleTrp 370
QY      359 GTTAGACAAACACAGTGGGGGCTGTACCATCGATCATCTCTGCGCAGGTACACCA 418
      :      :
Db     371 -----ProAlaGlnThrHIsPro-Pro----- 377
QY      419 GCAACAGGGGACACAGGAGAGCCAGTACAGAGCTTCCCAAGTTCACTGCGGAGAG 478
      :      :
Db     378 -GlnProHIsAArgProThrAArgProGlnLeu-----SerProGly----- 390
QY      479 GGCAGAAAGACGATCCGCTCATCGGAGGTGGTGTCTCTCTCATTTGCCCTGGTGTTCG 538
      :      :
Db     391 -----ThrSerLeuAlaAlaSerSerSerSerHIsTrpProSerSe 404
QY      539 CTATCATCTCTTCCAGTTCTGACAGGGCCACACAGGATACAGTACAGTACAGAGAG 598
      :      :
Db     404 rThrThrSerThrThrSerSerThrThrSerThrThrThrThrThrThrThrAArgAr 424
QY      599 GAGAGCTGTCCCAAGACAGCTGTGTGGCTGTGACCGGGGTGTGACGTCAAGCTGAAGAGT 658
      :      :
Db     424 gThrThrThrProThrThr-----ThrThrAArgAArgHrThrHrThrAAsnLysP 440
QY      659 GACGAGCTGGCGCTGAGAGGTTGACTGGGACAAAGTCTGCTTAAATCTACTCTGGG 718

```

```

Db      440 oThArgProtyrGlnArgProthrAlaThrSerSerSerSerthr----- 458
Qy      719 TCCTCCCATCATGTCGCTCCATCTGTAGCAGCACTGGAATGACTCTACTACAGAGAG 778
Db      459 -----ThSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 468
Qy      779 ACCTGACGAGCTGGGTTTCGAG-----AGTGTCACCGGAGAAC 819
Db      468 GPRIILSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 488
Qy      820 CGAGGTGCCGAGGAGATTGGCCAAAGCTTCTCAATCTTGAGATTAACAATCCACCAT 879
Db      488 rGlnProThrHisArgThrProValLeuAlaThrSerGlyLeuGlnThrSerSerSer 508
Qy      880 CCAGAAAGCTCCACAGGCTGAATGCCCTCCAG---CGATATCTCTCCACAGT 936
Db      508 rAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 528
Qy      937 TTCCACATGGGAGTGAAGGCCATGACC-----GGCGGAGTCGCGGAGGGGGCT 987
Db      528 gSerGlyCysGlyValProThrLeuAlaArgProGlnThrArgLeuAlaGlyGlySe 548
Qy      988 GGCTCTCGAATAGCAAGTGGCTTGGCAAGTGAATCTGCAC-----TT 1029
Db      548 rAlaAlaPheGlyArgTrpProTrpGlnValSerValArgArgThrSerPhePheGly 568
Qy      1030 CGGACCAACCCACATCTGTGAGGAGCGCTGATGACGCCAGTGGGTGCTACTGCCGC 1089
Db      568 eSerSerThrHisArgCysGlyGlyAlaLeuIleAsnGlnAsnTrpIleAlaThrAla 588
Qy      1090 CCACTGC-----TTCTTCGTCACCCGGGGAAGAGTCTCTGAGAGGCTGGAAGGTGA 1140
Db      588 yHisCysValAspAspLeuLeuIleSerGlnIleArgGlyLeuArgValGly-----Gly 606
Qy      1141 CGCGGGAGCAGCAGCACTGACACGATTGGCCAG---GCAAGCTCCATTCGCGAGATCAT 1197
Db      606 rAspPheSerHisValGlnGlnGlnLeuProTyrlleGlnArgGlyAlaAlaValSyl 626
Qy      1198 CATCAACAGCAATTAACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTC 1257
Db      626 lValHisProIleArgSerPheLeuThrTyrlleGlyValAspLeuAlaValSyl 646
Qy      1258 CAAGCCCTGACCTGTCGCTGATCCACCTGCTTGGCTCCCATGATGACAGAC 1317
Db      646 uGlnProLeuGlnPheAlaProHisValSerProIleCysLeuPro-----Gly 663
Qy      1318 CTTAGAGCTC-----AATGAGACCTGCTGATGACAGGCTTGGACAGACAG 1365
Db      663 rAspSerLeuLeuIleGlyMetAlaThr-----ValThrGlyTrpGlyArgLeu 681
Qy      1366 GGAGACAGATGACAAGACATCCCTCTCTCGGAGGTGACGATCATCTCATGACTT 1425
Db      681 rGlnGlyGlyThrLeuProSer---ValLeuGlnGlyValSerValProIleValSer 700
Qy      1426 CAAGAATGCAATGACTACTGTC-----TATGACAGTTCATCTTACCCCAAGAT 1476
Db      700 nAspAsnCylSerSerMetPheMetArgAlaGlyArgGlnGlnPheIleProAsp 720
Qy      1477 GATGTGTCGGGAGACCTGTCGGGGGAGAGATCCCTGACGAGGAGAGACGGGGGACC 1536
Db      720 eLeuGlyAlaGlyIleGlyTrpGlyGlnAspSerSerSerSerSerSerSerSerSer 740
Qy      1537 TCTTGTCTGTAG---CAGAACCAACCGTGTGATCTGGCAGGTGTCACAGCTGGGGC 1593
Db      740 oLeuGlnAlaIleAspSerGlnAspGlyArgPhePheLeuAlaGlyIleIleSer 760
Qy      1594 AGGTGTGGCCAGGAACAACCTGTGTGTATACCAACAATGACAGAAGTCTTCCCTG 1653
Db      760 eGlyCysAlaGlnAlaAsnLeuProGlyValCysThrArgIleSerSerSerSer 780
Qy      1654 GATT 1657

```

```

Db      780 pile 781
RESULT 13
KFPHI
N:Alternate names: antithrombophilic factor C; plasma thromboplastin antecedent
C/Species: Homo sapiens (man)
C/Date: 13-Aug-1986 #sequence revision 26-May-1994 #text_change 05-Oct-2004
R/Accession: A27431, A00920, A37940
R/Author: R., Davie, E.W.; Chung, D.W.
Biochemistry 26, 7221-7228, 1987
A/Title: Organization of the gene for human factor XI.
A/Reference number: A27431, MUID:88107663, PMID:2827746
A/Accession: A27431
A/Molecule type: DNA
A/Residues: 1-625 <ASA>
A/Cross-references: UNIPROT:P03951; UNIPARC:UPI000000DB87; GB:M18295
A/Note: the sequence shown follows the authors' translation
R/Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
Biochemistry 25, 2417-2424, 1986
A/Title: Amino acid sequence of human factor XI, a blood coagulation factor with four tar
A/Reference number: A00920, MUID:86243360, PMID:3636155
A/Accession: A00920
A/Molecule type: mRNA
A/Residues: 1-625 <FUJ>
A/Cross-references: UNIPARC:UPI000000DB87; GB:M13142; NID:G182832; PIDN:AA52487.1; PID:
R/Mueller, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2056-2060, 1991
A/Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
A/Reference number: A37940, MUID:91152017, PMID:1998667
A/Accession: A37940
A/Molecule type: Protein
A/Residues: 28-33/35-49, 'X', 51-55, 'X', 57-63/70-75, 'X', 77-79/107-109, 'X', 111-112/132-139,
'280-282, 'X', 284/285-297/313-316, 'X', 318-319/320-326, 'X', 328-330/347-349/373, 'X', 375;
A/Cross-references: UNIPARC:UPI00000172B4; UNIPARC:UPI00000172B25; UNIPARC:UPI00000172B26;
B2B; UNIPARC:UPI00000172B2C; UNIPARC:UPI00000172B2D; UNIPARC:UPI00000172B2E; UNIPARC:UPI000
10000172B34; UNIPARC:UPI00000172B35; UNIPARC:UPI00000172B36; UNIPARC:UPI00000172B37; UNIPARC
C/Comment: The proenzyme consists of two identical chains linked by one or more disulfide
he active site, and a heavy chain, which associates with high molecular weight (HMW) kind
C/Genetics:
A/Gene: GDB:F11
A/Cross-references: GDB:119891; OMIM:264900
A/Map position: 4q35-4q35
A/Intons: 19/1; 73/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 526
C/Function:
A/Description: catalyzes the proteolytic activation of coagulation factor IX
A/Pathway: blood coagulation intrinsic pathway
C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C/Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro
F.1-18/Domain: signal sequence #status predicted <SIG>
F.19-387/Product: coagulation factor XIa heavy chain #status experimental <HCH>
F.19-108/Domain: apple repeat <AP1>
F.109-198/Domain: apple repeat <AP2>
F.199-288/Domain: apple repeat <AP3>
F.290-379/Domain: apple repeat <AP4>
F.388-625/Product: coagulation factor XIa light chain #status experimental <LCH>
F.388-618/Domain: trypsin homology <TRY>
F.20-103, 514-581, 571-599/Disulfide bonds: #status predicted
F.29/Diulfide bonds: interchain #status experimental
F.146-76, 50-56, 110-193, 136-165, 140-146, 200-283, 226-255, 230-236, 291-374, 317-346, 321-327, 380
F.190-126, 353, 450/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.339/Disulfide bonds: interchain #status predicted
F.387-388/Cleavage site: Arg-Ile (coagulation factor XIa) #status experimental
F.431, 480, 575/Active site: His, Asp, Ser #status predicted
F.491/Binding site: carbohydrate (Asn) (covalent) #status experimental
Alignment Scores:
Pred. No.: 4, 54e-22 length: 625
Score: 532.00 matches: 138
Percent Similarity: 45.5% conservative: 62
Best Local Similarity: 31.4% mismatches: 136
Query Match: 16.0% indels: 104
DB: 1 gaps: 17

```

```

US-10-806-370-11 (1-1748) x KFNU1 (1-625)
Oy 609 CCAAGCAGCGCTGCTGAGCGGAGGAGCTGCAAGCTGAAGAGTGAAGAGCTG- 667
Db 206 ProanthraValPheAla--AspSerAnilLeAspSerValMetAlaProAspAlaPheV 225
Oy 668 -----GGCTGGTGAAGTTT-----G 683
Db 225 aLcYgGLyArgIleCysThrRhIShISpRoglCysLeuPhePheThrPheSerGing 245
Oy 684 ACTGGAGCAAGTCT-----CTGCTTAAATCTACTCTGGCTCTCC 725
Db 245 IutPrProLyGluSerGlnArgAnLeuCyLeuLeuLys-----ThrsArg 261
Oy 726 ATCAGTGCTTCCC-----ATCTGAGCAAGTGAATGACTCTACCTACAGAGA 776
Db 261 IuSerGlyLeuProSerThrArgIleLySeryLysAlaLeuSerGlyPheSerLeug 281
Oy 777 AGACCTGCAGCAGCTGGGTTTCAGAGTGTCTACCGGACCAAGGAGTTGCCAGAGG 836
Db 281 InsSerCyArg-----HisSerIleProValPheCysHisSers 294
Oy 837 ATTTGCCAAGCTTCTCAATCTTGAGATCAACTCCACATCCAGAGAAAGCTCCACA 896
Db 294 erPheTyRhISAspThrAspPheLeuGlyGluLeuAspIleValAlaIaLysSerH 314
Oy 897 GGTCTGAATGC-----C 908
Db 314 IeGluAlaCyGlnIleLeuCySthrAnAlaValArgCyGlnPhePheThrTyThrP 334
Oy 909 CTTCACGAGCG-----TATATCTCCCTCAGAGTTTCCC 941
Db 334 roAlaGlnAlaSerCyAnGlnGlyLysCysTyLeuLeuLysSerSerAnG 354
Oy 942 ACTGGGAGCTAGAGGCGCATGACCGG----- 967
Db 354 lySerProThrLyLeuLhISglYArgIlyIleSerglyTyThrLeuArgLeuC 374
Oy 968 -----CGGATCGTGGAGAGGGGCGTGG 989
Db 374 yelYMetAspAnGluCyThrThyLySleLySProArgIleValISglYThrAlaS 394
Oy 990 CCTCGATAGCAGTGGCTTGGCAGTGAAGTGTGCTTGGCAGCAGC----- 1039
Db 394 erValArgGlyGluTrpTrpTrpGlnValThrlhIS-----ThrlhSerProThrg 412
Oy 1040 -----CACATCTGTGAGAGCAGCTCATTAAGCCCAAGTGGTGTCTACCTGCCGCCACT 1094
Db 412 IuArgHISleuCySglYglYSerIleIleGlyAsnGlnTrpIleLeuThrlaAlaHIS 432
Oy 1095 GCTTCTTC---GTGACCGCGGAGAGGTCCTGAGAGGCTGAGAGGCGTGAAGCGGGGACCA 1151
Db 432 yelPheTyGlyValGluSerProLySleLys-----ArgAlaTySerglyIleL 449
Oy 1152 GCAACCTGCACCAAGTGGCTGAGGAGCGCTCC-----ATTGCCAGATCATCATCA 1202
Db 449 euAnGlnInsSerGluIleLyGlnuAspThrSerPhePheGlyValGlnIuIleIleH 469
Oy 1203 ACAGCAATTACCGATAGAGACAGCAGCTAGATGCGCTTACCGGCTGTCCAAAGC 1262
Db 469 ISAspGlnTyLyMetAlaGluSerGlyTyAspIleAlaLeuLeuLySleuGluThrT 489
Oy 1263 CCTGACCGCTCGCATCATCCAGCTGCTGGTGGCTCCCATGATGAGACAGACCTTTA 1322
Db 489 hrValInsTyThrAspSerGlnArgProIleCySLeuProSerLySglYAspArgAnV 509
Oy 1323 GCCTCAATAGACCTGTGATCAAGGCTTGGCAGAGCAGGAGAGCAGATGACAGAGA 1382
Db 509 alIleTyThrAspCyTrpValThrglyTyTrgLy---TyArgLySleuArgAspLySI 528
Oy 1383 CATCCCTTCTCGGGAGGTGACAGGTCAATTCATGAGTCAAGAAATGCAATGACT 1442

```

```

Db 528 IeGlnAnThrLeuGlnAlaLySleIleProLeuValThrAsnGluGluCySgluLyS 548
Oy 1443 ACTTGCTATGACAGTTAC---CTTACCCCAAGATGATGTGTGCTGGGACCTTGGT 1499
Db 548 rg-----TyArgGlyHISleLySleIleThrlhISyMetIleCyAlaISglYTrArgGlu 566
Oy 1500 GGGGAGAGACTCTGCCAGGAGACAGACGGGGGCGCTTTGTCTGTGAGCAGAAACACC 1559
Db 566 lyGlyLysAspAlaCyLeuGlyAspSerGlyGlyPProLeuSerCyLeuHISAnGluV 586
Oy 1560 GCTGTACTCTGGCAGGTGTACCCAGCTGGGGCAGACAGCTGTGGCCAGAGAAACAACTG 1619
Db 586 alTrpHISleuValISglYIleThrserTrpGlyGluGlyAlaGlnArgIuArgProG 606
Oy 1620 GTGTGTACCAACCAAGTCTCTCCCTGATTTACAGCAGATGAGAGAC 1675
Db 606 lyValTyThrAnValAlaGluTyValAlaAspTrpIleLeuGluLySleThrlhAla 624

RESULT 14
PLBO
plasma (EC 3.4.21.7) precursor - bovine
N.Alternate names: plasminogen
C.Species: Bos primigenius taurus (cattle)
C.Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C.Accession: S45046; A25835; I45961; S03736
R.Berglund, L.; Andersen, M.D.; Petersen, T.E.
Submitted to the EMBL Data Library, May 1994
A.Description: Cloning and characterization of the bovine plasminogen cDNA.
A.Reference number: S45046
A.Accession: S45046
A.Molecule type: mRNA
A.Residues: 1-812 <BER>
A.Cross-references: UNIPROT:P06868; UNIPARC:UPI0000043BEV; EMBL:X79402; NID:9494962; PID:
A.Experimental source: liver
A.Note: it is uncertain whether Met-1 or Met-8 is the initiator
R.Schaller, U.; Moser, P.W.; Dannegger-Mueller, G.A.K.; Rosselet, S.J.; Kamper, U.; Rickl
Bur. J. Biochem. 149, 267-278, 1985
A.Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasmi
A.Reference number: A25835; MUID:85203906; PMID:3846532
A.Accession: A25835
A.Molecule type: protein
A.Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
A.Cross-references: UNIPARC:UPI000016C365; GB:K02935; NID:9163551; PID:AAA30714.1; PID:
R.Malinowski, D.P.; Sadler, J.E.; Davies, E.W.
Biochemistry 23, 4243-4250, 1984
A.Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A.Reference number: I45961; MUID:85023311; PMID:6148961
A.Accession: I45961
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 706-743, 'R', 745-812 <MAL>
A.Cross-references: UNIPARC:UPI000016C365; GB:K02935; NID:9163551; PID:AAA30714.1; PID:
R.Brunholz, R.A.; Lerch, P.G.; Schaller, U.; Rickli, E.E.; Lergler, W.; Manneberg, M.;
Bur. J. Biochem. 114, 465-470, 1981
A.Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A.Reference number: S03735; MUID:81212097; PMID:7238497
A.Accession: S03736
A.Molecule type: protein
A.Residues: 27-83 <BRU>
A.Cross-references: UNIPARC:UPI0000172B96
C.Function:
A.Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
in the walls of the graafian follicle; also activates the urokinase-type plasminogen acti
A.Pathway: fibrinolysis
C.Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology;
C.Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma;
F:1-66/Domain: signal sequence #status predicted <SIG>
F:8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F:27-812/Product: plasminogen #status experimental <PRO>
F:27-103/Domain: activation peptide #status experimental <APT>
F:104-583, 584-812/Product: plasmin chain A #status experimental <MAT>
F:110-188/Domain: kringle homology <KRI>

```


A/Title: Molecular cloning, tissue-specific expression, and cellular localization of human
 A/Reference number: A57014; MUID:95286644; PMID:7768952
 A/Accession: A57014
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-343 <RBS>
 A/Cross-references: UNIPROT:Q16651; UNIPARC:UPI0000046DCB; GB:L41351; NID:9862304; PIDN:
 A/Experimental source: prostate
 A/Note: parts of this sequence were determined by protein sequencing
 R/Yu, J.X.; Chao, L.; Chao, J.
 J. Biol. Chem. 269, 18843-18848, 1994
 A/Title: Prostatein is a novel human serine proteinase from seminal fluid. Purification,
 A/Reference number: A54866; MUID:94308140; PMID:8034638
 A/Accession: A54866
 A/Molecule type: protein
 A/Residues: 45-64 <YUN>
 A/Cross-references: UNIPARC:UPI0000172AFB
 C/Genetic8:
 A/Gene: GDB:PRSS8
 A/Cross-references: GDB:676446; OMIM:600823
 A/Map position: 16p11.2-16p11.2
 C/Superfamily: trypsin; trypsin homology
 C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein
 F/1-32/Domain: signal sequence #status predicted <SIG>
 F/33-44/45-343/Product: prostatein #status predicted <MAT>
 F/33-44/Domain: prostatein light chain #status predicted <CHL>
 F/45-343/Domain: prostatein heavy chain #status predicted <CHN>
 F/45-281/Domain: trypsin homology <TRY>
 F/323-341/Domain: transmembrane #status predicted <TM1>
 F/37-154,70-86,166-244,201-223,234-262/Disulfide bonds: #status predicted
 F/85,134,238/Active site: His, Asp, Ser #status predicted
 F/159/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 7,04e-21 | Length: | 343 |
| Score: | 510.50 | Matches: | 103 |
| Percent Similarity: | 57.5% | Conservative: | 42 |
| Best Local Similarity: | 40.9% | Mismatches: | 94 |
| Query Match: | 15.3% | Indels: | 13 |
| DB: | 1 | Gaps: | 5 |

US-10-806-370-11 (1-1748) x A57014 (1-343)

```

QY 944 TCGGAGCTGAGGAGCCATGACCGGAGGATCGTGGAGGAGGCGCTGGCCTCGAATGACAAG 1003
Db 37 CysGlyVal--AlaProGlnAlaArgLleThrGlyGlySerSerAlaValAlaGlyGln 55
QY 1004 TGGCCTTGGCAAGTGAAGTGTGACCTTGGCAGCACCAACCATCTGTGGAGGACGCTCAT 1063
Db 56 TrpProTrrGlnValSerLleThrTyrGlnGlyValHisValCysGlyGlySerLeuVal 75
QY 1064 GACGCCCAAGTGGAGTGTCTACTACCGCCCACTGCTTCTTGGACCCGGAGAGAAGTCTCTG 1123
Db 76 SerGlnGlnTrpValLeuSerAlaHisCysPheProSerGlnHisIleLys----- 93
QY 1124 GAGGCTGGAAGGTGTACGCGGAGCACCAAGCAACTGCAACGATTGCTGAGGACGCTCC 1183
Db 94 GlnAlaTyrGlnValValLysGlnGlyAlaHisGlnLeuAspSerTyrSerGlnAspAlaLys 113
QY 1184 ATTGCC-----GAGATCATCATCAACGCAATTACACCGATGAGGAGAGACGACTAT 1234
Db 114 ValSerThrLeuLysAspLleLleProHisProSerTyrLeuGlnGlnGlySerGlnGly 133
QY 1235 GACATCGCCCTCATGCGGCTGTCCAGGCGCTGACCCGCTGACCATCCACCTGCT 1294
Db 134 AspLleAlaLeuLeuGlnLeuSerAlaGrrProLleThrPheSerArgTyrLleArgProLle 153
QY 1295 TGCCTCCCATGATGATGACAGACCTTTAGCTCAATGAGACCTGCTGATCAGAGCTTT 1354
Db 154 CysLeuProAlaAlaAsnAlaSerPheProAsnGlyLeuHisCysThrValThrGlyTyr 173
QY 1355 GGCAGACACGAGGAGACAGATGACACAGATCCCTTC---CTCCGGAAGTGCAGGTC 1411

```

```

Db 174 GlyHisValAlaProSerValSerLeuLeuThrProLysProLeuGlnGlnLeuGlnVal 193
QY 1412 AATCTCATTCGACTTCAAGAAATGCATGACTGTGCTATGACAGT----- 1459
Db 194 ProLeuIleSerArgGlnThrCysAsnCysLeuTyrAsnLleAspAlaLysProGlnGln 213
QY 1460 -----TACCTTACCCCAAGATGATGTGTCTGGGAGACCTTCGTGGGGCGAGACTCC 1513
Db 214 ProHisPheValGlnGlnAspMetValCysAlaGlyTyrValGlnGlyGlyLysAspAla 233
QY 1514 TGCAGGAGACAGACGGGGGCGCTTGTCTGTGTAGACAGAAACCGCTGTACTCTGCA 1573
Db 234 CysGlnGlyAspSerGlyGlyProLeuSerCysProValGlnGlyLeuTrrPrrLeuThr 253
QY 1574 GGTGTACCAAGCTGGGCGACAGGCTGTGGCCAGAGAAACAACTGGTGTGTACCCAAA 1633
Db 254 GlyLleValSerTrpGlyAspAlaCysGlyAlaArgAsnArgProGlyValTyrThrLeu 273
QY 1634 GTGACAGAAGTCTTCCCTGGATTTCACGACAAGTG 1669
Db 274 AlaSerSerTyrAlaSerTrpLleGlnSerLysVal 285

```

Search completed: September 16, 2006, 01:54:45
 Job time : 107.5 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rnp**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 01:24:21; Search time 106.6 Seconds
(without alignments)
4350.452 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 3333
Sequence: 1 ctccgagacatcgagagag9.....ggcgcgtgactcgagaaa 1748

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 5699196

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWB/spool/US10806370/runat.15092006.105625.9965/app.query.fasta-1
-DB=UniProt -QEXT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=humana0.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=ds802p
-USER=US10806370 @CGN 1 1 891 @runat.15092006.105625.9965 -NCPU=6 -ICPU=3
-NO MMAP -NCG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

UniProt_7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------|
| 1 | 2956 | 88.7 | 581 | 1 | TMPSD_HUMAN |
| 2 | 2510.5 | 75.3 | 543 | 1 | TMPSD_MOUSE |
| 3 | 891 | 26.7 | 359 | 2 | Q4SPG0_TETNG |
| 4 | 784.5 | 23.5 | 484 | 2 | Q311V4_PANPA |
| 5 | 783.5 | 23.5 | 484 | 2 | Q311V3_PPRIM |
| 6 | 780.5 | 23.4 | 490 | 1 | Q311V5_PANTR |
| 7 | 777.5 | 23.3 | 484 | 1 | TMPS2_MOUSE |
| 8 | 774.5 | 23.2 | 484 | 2 | Q311U4_MACMU |
| 9 | 773 | 23.2 | 490 | 2 | Q6P7D7_RAT |
| 10 | 771.5 | 23.1 | 490 | 2 | Q3UKK3_MOUSE |
| 11 | 770.5 | 23.1 | 490 | 2 | Q7NN04_MOUSE |
| 12 | 768.5 | 23.1 | 492 | 2 | Q6GKT7_HUMAN |
| 13 | 767.5 | 23.0 | 484 | 2 | Q311U8_HYUSY |
| 14 | 766.5 | 23.0 | 492 | 1 | TMPS2_HUMAN |
| 15 | 766.5 | 23.0 | 492 | 2 | Q96T73_HUMAN |
| 16 | 764 | 22.9 | 490 | 2 | Q920K3_RAT |

| | | | | | |
|----|-------|------|------|---|--------------|
| 17 | 763.5 | 22.9 | 453 | 2 | Q812A6_MOUSE |
| 18 | 763.5 | 22.9 | 453 | 2 | Q3T206_MOUSE |
| 19 | 762.5 | 22.9 | 475 | 1 | TMPS3_MOUSE |
| 20 | 762.5 | 22.9 | 453 | 2 | Q2M1G4_MOUSE |
| 21 | 759.5 | 22.8 | 486 | 2 | Q5PR46_BRARE |
| 22 | 736.5 | 22.1 | 538 | 2 | Q5USC7_HUMAN |
| 23 | 731 | 21.9 | 454 | 1 | TMPS3_HUMAN |
| 24 | 713.5 | 21.4 | 722 | 2 | Q6NUP5_XENLA |
| 25 | 713 | 21.4 | 767 | 2 | Q9DCR2_XENLA |
| 26 | 697.5 | 20.9 | 437 | 1 | TMPS4_HUMAN |
| 27 | 696 | 20.9 | 445 | 2 | Q8C1J7_RAT |
| 28 | 689.5 | 20.7 | 439 | 2 | Q5RDY7_PONPY |
| 29 | 684.5 | 20.5 | 435 | 1 | TMPS4_MOUSE |
| 30 | 684 | 20.5 | 388 | 2 | Q4RRR7_TETNG |
| 31 | 668.5 | 20.1 | 371 | 2 | Q8C1J6_RAT |
| 32 | 664 | 19.9 | 455 | 1 | TMPS5_MOUSE |
| 33 | 659 | 19.8 | 455 | 2 | Q8CDR0_MOUSE |
| 34 | 644.5 | 19.3 | 457 | 1 | TMPS5_HUMAN |
| 35 | 620.5 | 18.6 | 445 | 2 | Q3U0U6_MOUSE |
| 36 | 612.5 | 18.4 | 436 | 1 | HEPS_MOUSE |
| 37 | 611.5 | 18.3 | 417 | 1 | HEPS_HUMAN |
| 38 | 609.5 | 18.3 | 417 | 2 | Q5R5E8_PONPY |
| 39 | 600 | 18.0 | 326 | 2 | Q7Z2B0_BRARE |
| 40 | 592.5 | 17.8 | 416 | 1 | HEPS_RAT |
| 41 | 589.5 | 17.7 | 572 | 1 | TMPS7_MOUSE |
| 42 | 585.5 | 17.6 | 829 | 2 | Q4PCJ3_MOUSE |
| 43 | 582.5 | 17.5 | 730 | 2 | Q4RH07_TETNG |
| 44 | 580 | 17.4 | 1019 | 1 | ENTK_HUMAN |
| 45 | 580 | 17.4 | 1019 | 2 | Q2NKL7_HUMAN |

ALIGNMENTS

RESULT 1
TMPSD_HUMAN STANDARD; PRT; 581 AA.
ID Q9BYE2; Q86YM4; Q96JY8; Q9BYE1;
AC 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT 15-MAR-2005, sequence version 2.
DT 07-MAR-2006, entry version 29.
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
protease) (Membrane-type mosaic serine protease).
GN Name=TMPSR13; Synonyms=MSP, TMPSR11;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [mRNA] (ISOFORMS 1 AND 3), AND TISSUE SPECIFICITY.
RC TISSUE=Lung;
RX MEDLINE=21167393; PubMed=11267681; DOI=10.1016/S0167-4781(01)00184-1;
RA Kim D.R., Sharmin S., Inoue M., Kido H.,
RT "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung."
RL Biochim. Biophys. Acta 1518:204-209(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [mRNA] (ISOFORM 2).
RA Park T.J., Park W.J.,
RT "Homo sapiens transmembrane protease, serine 6 (TMPSR56) mRNA."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE mRNA] (ISOFORM 4).
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakematsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nishitani K.,
RA Murakami K., Yasuda T., Watanabe T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,

RA Ninoiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hozaka S., Chiba Y., Ichida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida S., Hotta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togaya S., Komai T., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsumura H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hisigaki H., Watanabe T., Sugiyama A., Takenoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togeashi T., Oyama M., Heta H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RT Nat. Genet. 36:40-45(2004).
RL [4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP Human chromosome 11 international sequencing consortium;
RG Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUPPLEMENTARY LOCATION: Membrane; single-pass type II membrane
CC protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=WSP, Large form; Membrane-type;
CC IsoId=Q9BYE2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9BYE2-2; Sequence=VSP_013103, VSP_013104;
CC Note=No experimental confirmation available;
CC Name=3; Synonyms=WSP, Small form;
CC IsoId=Q9BYE2-3; Sequence=VSP_013099, VSP_013102;
CC Name=4;
CC IsoId=Q9BYE2-4; Sequence=VSP_013100, VSP_013101;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Isoform 1 and isoform 3 are predominantly
CC expressed in lung, placenta, pancreas, and prostate. Isoform 3 is
CC weakly expressed in testis and peripheral blood lymphocytes.
CC -1- POLYMORPHISM: The repeat A-S-P-A-[GLQR] is polymorphic and the
CC number of copies varies between 12 to 14.
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
CC -1- SIMILARITY: Contains 1 SRCR domain.
CC -1- CAUTION: Ref.2 has referred to this protein as TMPRSS6.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
CC EMBL; AB048796; BAB39741.1; -; mRNA.
CC EMBL; AB048797; BAB39742.1; -; mRNA.
CC EMBL; AY190317; AAO38062.1; -; mRNA.
CC EMBL; AK027798; BAB55376.1; -; mRNA.
CC EMBL; AF002962; -; NOT_ANNOTATED_CDS; Genomic DNA.
CC HSSP; P00760; IEZX.
CC DR MEROPS; S01.087; -.
CC DR Ensembl; ENSG00000137747; Homo sapiens.
CC DR HGNC; HGNC:29808; TMPRSS13.
CC DR GO; GO:0016021; C-integral to membrane; NAS.
CC DR GO; GO:0004552; F-actin-type endopeptidase activity; NAS.
CC DR GO; GO:000508; P-peptidolysis and peptidolysis; NAS.
CC DR InterPro; IPR002172; LDL_rcpt_A.
CC DR InterPro; IPR001254; Peptidase_S1_S6.
CC DR InterPro; IPR001314; Peptidase_S1A.
CC DR InterPro; IPR001190; Src_rcpt.
CC Pfam; PF00057; Ldl_rcpt_a; 1.
CC Pfam; PF00089; Trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_spec; 1.
DR PROSITE; PS01209; LDLRA_1; FALSE_NEG.
DR PROSITE; PS00069; LDLRA_2; FALSE_NEG.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS00287; SRCR_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Alternative splicing; Glycoprotein; Hydrolase; Membrane; Polymorphism;
KW Protease; Repeat; Serine protease; Signal-anchor; Transmembrane.
FT CHAIN 1 581
FT FT
FT TOPO_DOM 1 160
FT FT
FT TRANSMEM 161 181
FT FT
FT TOPO_DOM 182 581
FT FT
FT REPEAT 9 13
FT FT
FT REPEAT 14 18
FT FT
FT REPEAT 19 23
FT FT
FT REPEAT 24 28
FT FT
FT REPEAT 29 33
FT FT
FT REPEAT 34 38
FT FT
FT REPEAT 39 43
FT FT
FT REPEAT 44 48
FT FT
FT REPEAT 49 53
FT FT
FT REPEAT 54 58
FT FT
FT REPEAT 59 63
FT FT
FT REPEAT 64 68
FT FT
FT REPEAT 69 73
FT FT
FT REPEAT 74 78
FT FT
FT REPEAT 79 83
FT FT
FT REPEAT 84 88
FT FT
FT DOMAIN 190 320
FT FT
FT DOMAIN 199 221
FT FT
FT REGION 321 554
FT FT
FT REGION 9 88
FT FT
FT REGION 14 68
FT FT
FT COMBIAS 9 89
FT FT
FT ACT_SITE 361 361
FT FT
FT ACT_SITE 409 409
FT FT
FT ACT_SITE 506 506
FT FT
FT CARBOHYD 250 250
FT FT
FT CARBOHYD 287 287
FT FT
FT CARBOHYD 400 400
FT FT
FT CARBOHYD 440 440
FT FT
FT DISULFID 245 309
FT FT
FT DISULFID 258 312
FT FT
FT DISULFID 346 362
FT FT
FT DISULFID 443 512
FT FT
FT DISULFID 475 491
FT FT
FT DISULFID 502 530
FT FT
FT VARSPPLIC 146 181
FT FT
FT FT
FT VARSPPLIC 423 486
FT FT
FT FT
FT VARSPPLIC 555 581
FT FT
FT FT
FT VARSPPLIC 487 581
FT FT
FT FT
FT VARSPPLIC 555 581
FT FT
FT FT
FT VARSPPLIC 555 581
FT FT
FT FT
FT VARSPPLIC 555 581
FT FT
FT FT
FT VARSPPLIC 559 581
FT FT
FT FT
FT CONFLICT 187 187
FT FT
FT CONFLICT 192 192
FT FT
FT CONFLICT 201 201
FT FT
FT CONFLICT 206 206

AHRPACIPMGQFISINETCWTGFKTRETDTKSPFLR
EVQVGLIDFKKCNLYVDSYLT -> GEGICTRSPAPQ
OHPLQSHLSASVNSYRPPKASQKSTLKDPMHRCFI
IRRTREAGL (in isoform 4).
/FTId=VSP_013100.
Missing (in isoform 4).
/FTId=VSP_013101.
VRSIQDTAPSLRLTSGSGDPCGAPRV -> SEVFRKRS
(in isoform 3).
/FTId=VSP_013102.
VRSI -> SSAG (in isoform 2).
/FTId=VSP_013103.
Missing (in isoform 2).
/FTId=VSP_013104.
H -> Y (in Ref. 2).
Y -> H (in Ref. 1); BAB39741).
K -> E (in Ref. 2).
C -> R (in Ref. 2).

FT CONFLICT 259 259 Q -> R (in Ref. 1).
 FT CONFLICT 298 298 E -> H (in Ref. 1).
 FT CONFLICT 496 496 H -> R (in Ref. 2).
 SQ SEQUENCE 581 AA; 62680 MM; A4935CCCP3D29EE CRC64;

Alignment Scores:

Pred. No.: 3.66e-158 Length: 581
 Score: 2956.00 Matches: 553
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 88.7% Indels: 0
 DB: 1 Gaps: 0

US-10-806-370-11 (1-1748) x TMPSD_HUMAN (1-581)

QY 11 ATGAGAGGAGACACCGAGATGATCTTCACGAAACAACCTTAGAGCATCT 70
 Db 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAAGCCAGGCATCTCAGCTGGGACACCTCCAGCGGGGATCTCCAGCCAGGCATCT 130
 Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAAGCCAGGCATCTCAGCTGGGACACCTCCAGCGGGGATCTCCAGCCAGGCATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAAGCTGTACACCTCCAGCGGGGATCTTCAGCGCGGGATCTCCAGCCAGGCATCT 250
 Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAAGCCCGGCATCTCCGGCTTGGGATCACTTCCAGGTCCTCATCCGGCAGGTCATCA 310
 Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerLeuSerSerSerGlyArgSerSer 100
 QY 311 TCCGSCAGCTCAGCGCTGGTGAACAACCTCCCAACAGAGTGTCTGTAGAGACA 370
 Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
 QY 371 CCAAGTGGGGCTGTATCCATCCATCATCTCTCCAGGTGACAGCAACAGAGGCC 430
 Db 121 ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGCCAGGTGACAGCGCTGCCAAGTTCACTCCGGGAGGGGACAGACAG 490
 Db 141 ThrArgIleuSerProGlyThrSerLeuProIleThrArgGlyGlnGlyGln 160
 QY 491 CTACCGCTCATCGGGGCTGCTGCTCTCTCATCTCCCTGGGTGTTGCTCATCATCTC 550
 Db 161 LeuProIleuIleGlyCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIleLeu 180
 QY 551 TTCAGATTCTGGCAGGGGACACAGAGGATCAGTACAGAGACAGAGAGAGCTGTCCC 610
 Db 181 PheGlnIleThrArgGlnGlnIleThrGlyIleArgTyrIleGlnIleArgGlnSerCysPro 200
 QY 611 AAGGACCTGTCTTCCCTGTACGGGGGTGGATCTGCAAGCTGAAGAGTACAGACTGGCC 670
 Db 201 LysHisAlaValAlaArgCysAspGlyValAlaAspCysIleuSerSerAspGlnIleGly 220
 QY 671 TGGGTGAGGTTTGAAGTGAAGTCTGCTTAAATCTACTCTGGAGTCTCCCATCAG 730
 Db 221 CysValArgPheAspIlePheAspIleSerLeuLeuIleTyrSerGlySerSerHisGln 240
 QY 731 TGGCTTCCATCTGTAGACAGCAACTGGAATGACTCTTACTAGAGAAAGCTGCGACAG 790
 Db 241 TrpLeuProIleCysSerSerAsnTrpAsnAspSerTyrSerGlnIleCysGlnGln 260
 QY 791 CTGGGTTTCAGAGATGCTCACCGGACCAACGAGGTGCCACAGGATTTTGGCCAAAGC 850
 Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGlnValAlaHisArgAspPheAlaAspSer 280
 QY 851 TTCTCAATCTTGAGATCAACTCCACCATCCAGGAAAGCTTCACAGGTCTGAATGCCCT 910

Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnIleuSerLeuHisArgSerGluCysPro 300
 QY 911 TCCAGAGGATATATCTCCCTCCAGTGTCTCCACTGGCGAGTACAGGAGCATGACCGGCGG 970
 Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
 QY 971 ATCGTGGAGGGGCGTGGCGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1030
 Db 321 IleValGlyIleAlaLeuAlaSerAspSerTyrTrpProIleGlnValSerLeuHisPhe 340
 QY 1031 GGCACCAACCAATCTGTGGAGGACCGCTTATGACGCCCAAGTGGGTCTCACTGCGGCC 1090
 Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
 QY 1091 CACTGCTTCTCGTGAACCGGGAGAAAGTCTCTGGAGGCTGGAAGGTGTACCGGGGACC 1150
 Db 361 HisCysPhePheValThrArgGlnIleValLeuGlnIleTyrIlePheValTyrAlaGlyThr 380
 QY 1151 AGCAACCTTGACACAGTTGCTTGAAGGACCTCCATTGCCGATCATCATCAACAGCAAT 1210
 Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleAsnSerAsn 400
 QY 1211 TACACGATGAGAGAGACGACTATGACATGCGCTTCAATGCGGCTGTCCAAAGCCCTGACC 1270
 Db 401 TyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArgLeuSerIleProLeuThr 420
 QY 1271 CTGTCCGCTCATCATCCACCTGCTGGCTCCCATGATGAGACGACCTTAGGCTCAAT 1330
 Db 421 LeuSerIleHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
 QY 1331 GAGACCTGTGATATCACAGGCTTTGGCAAGACAGGAGACAGATGACAAAGACATCCCC 1390
 Db 441 GluThrCysTrpIleThrArgIlePheGlyIleThrArgGlnThrAspAspIleThrSerPro 460
 QY 1391 TTCCTCCGGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGACTTGGTC 1450
 Db 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheIleCysAsnAspTyrLeuVal 480
 QY 1451 TATGACAGTTACCTTACCAGAGATGATGTGTCTGGGACCTTCCGTGGGGGACAGAGAC 1510
 Db 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuHisGlyGlyArgAsp 500
 QY 1511 TCTGTCAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGACAGAACCGCTGTACTTG 1570
 Db 501 SerCysGlnIleAspSerGlyGlyProLeuValCysGlnGlnAsnAlaArgTrpTyrLeu 520
 QY 1571 GCAAGTGTACACAGTGGGACAGCGCTGTGGCCAGAAACAAACCTGGTGTATCAC 1630
 Db 521 AlaGlyValIleThrSerTrpGlyThrGlyCysGlyGlnArgAsnIleProGlyValTyrThr 540
 QY 1631 AAGGTGACAGAAGTCTTCCCTGATTTACAGCAAGTTGAG 1672
 Db 541 LysValThrGlnValLeuProTrpIleTyrSerIleuMetGln 554

RESULT 2

TMPSD_MOUSE STANDARD; PRT; 543 AA.
 AC Q5U405; Q8CFE0; Q91VQ8;
 DT 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
 DT 15-MAR-2005, sequence version 2.
 DT 07-MAR-2006, entry version 16.
 DE Transmembrane protease, serine 13 (BC 3.4.21.-) (Mosaic serine
 protease) (Membrane-type mosaic serine protease).
 GN Name=tmprsb13; Synonym=msp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=B5/BGFP; and FVB/N;
 TISSUE=Mammary tumor, and trophoblast stem cell;

Db 162 PhetrpethrparcgllyhsthrcllyllyllystlyllysgluProleuGluSerCysPro 181
 QY 611 AAGCAACCTGTTGCTGCTGACGGGGGTGGTGGACCTGCAAGAGTGAAGAGTGGG 670
 Db 182 ILehslalvalArgCysaspGlyvalValaspCyslysmetLysSeraspGluLeuGly 201
 QY 671 TGGCTGAGTTGCTGCTGCAACCTGCTGTTAAATCTACTGCTGGGCTCTCCCATCAG 730
 Db 202 CysValArgPheaspPripaspLysSerleuLeuLysValTyrSerGlySerSerGlyGlu 221
 QY 731 TGGCTTCCACTGCTGACAGCAAGCACTGGAATGACTCTACTGACAGAAAGCTGGCAG 790
 Db 222 TrrpneurolCysSerSerSerSerTrrpAspPheaspPripaspLysSerGlyGlu 241
 QY 791 CTGGGTTTCCAGAGTGTCTACCCGCAACCAAGAGTGGTCCCAAGGATTTTGGCAACAG 850
 Db 242 LeucllyPheaspSerAlaTyrArgThrThrcllylValAlahlsAyspPripLethSerSer 261
 QY 851 TTCTCATCTTGAGATCAACCTCCACCATCCAGGAAAGCTCCACAGGTCTGAATGCCCT 910
 Db 262 PheleuSerGluTyrAspThrThrcllylLeuGluSerleuTyrArgSerGluCysSer 281
 QY 911 TCCAGCGGTATATCTCCCTCCAGTGTCTCCACTGCGACTGAGGCGCATGACCGGCGG 970
 Db 282 SerArgArgTyrValSerleuGluCysSerthlCysgllyLeuArgAlaMetThrcllyArg 301
 QY 971 ATCTGGAGGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGTCTGCACTTC 1030
 Db 302 ILevalGlyAlaLeuThrSerGluSerLysTrrpTrrpGluValSerleuHlsPhe 321
 QY 1031 GGCACCAACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1090
 Db 322 GlyThrThrthrlstleCysgllyGlyThrleuLeaspAlaglnThrValLeuThrAlaAla 341
 QY 1091 CACTGCTTCTTCTGACCCCGGAGAGAGTCTCTGAGGAGGCTGGAAGGTAGCGCGGAGCC 1150
 Db 342 HlsCysPhePheValThrArgGluLysleuLeucllyGlyTrrpLysValTyrAlaGlyThr 361
 QY 1151 AGCAACCTGCAACGAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1210
 Db 362 SerAsnLeuHlsGlnLeuProGluAlaAlaSerleSerGlnlleleleAsnGlyAsn 381
 QY 1211 TACACCGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
 Db 382 TyrThrAspGlnGlnAspPripaspLysLeuAlaLeuLethAspLysSerProleuThr 401
 QY 1271 CTGTCCTCCCTCAATCAACCTGCTGCTGCTCCCATGAGTGGAGAGGAGGAGGAGGAG 1330
 Db 402 LeuSerLethHlsleHlsPripaspLysleuProMetHlsGlyGlnThrPheGlyLeuAsn 421
 QY 1331 GAGACCTGCTGATCAACGAGGCTTTGGCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1390
 Db 422 GluThrCysTrrpLethGlnGlyPheGlyLysThrLysGlnThrAspGluLysThrSerPro 441
 QY 1391 TTCTCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450
 Db 442 PheleuArgGluValGlnValAsnLeuLeaspPheLysLysCysAspAspTrrpLeuVal 461
 QY 1451 TATACAGATTACTTACCCCAAGAGTATGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAG 1510
 Db 462 TyrAspSerTrrpLeuThrProArgMetCysAlaGlyAspLeuAspGlyGlyArgAsp 481
 QY 1511 TCTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1570
 Db 482 SerCysGlnGlyAspSerGlyGlyProleuValCysGlnGlnAsnAsnArgTrrpTrrpLeu 501
 QY 1571 GCAGGTGTCACAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1630
 Db 502 AlaGlyValThrSerTrrpGlyThrGlyCysGlyGlnLysAsnLysPripGlyValLysThr 521
 QY 1631 AAGGTGACAGAGTCTTCTCTGATTTTACAGCAAGATGAGGAGGAGGAGGAGGAGGAGGAG 1690
 Db 522 LysValThrGluValLeuProTrrpLethTrrpArgLysMetGluSerGluValArgPheArg 541

QY 1691 AATGCC 1696
 Db 542 LysSer 543
 RESULT 3
 ID QASPGO_TETNG PRELIMINARY; PRT; 359 AA.
 AC QASPGO; 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DT 07-FEB-2006, entry version 4.
 DE Chromosome 16 SCAP14537, whole genome shotgun sequence. (Fragment).
 OS Oryzias latipes (Green puffer).
 GN Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphia; Acanthopterygii; Percomorphia; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RX PubMed=15496914; DOI=10.1038/nature03025;
 RA Jallou O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
 Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 Micaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
 Daeilva C., Salenouat M., Levy M., Boudet N., Castellano S.,
 Anthouard V., Ubhi C., Castell V., Katinka M., Vacherie B.,
 Biemont C., Skalli Z., Cactolico L., Poullain J., De Berardinis V.,
 Cruaud C., Duprat S., Broctier P., Coutanceau J.-P., Gouzy J.,
 Parra G., Lardier G., Chappell C., McKernan K.J., McEwan P., Bosak S.,
 Kellis M., Wolfe J.-N., Guigo R., Zody M.C., Mesirov J.,
 Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 Lander V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype."
 RL Nature 431:946-957(2004).
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope, Whitehead Institute Centre for Genome Research;
 Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC
 EMBL: CAPE01014537; CAF97472.1; -; Genomic_DNA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005044; F:scavenger receptor activity; IEA.
 DR GO: GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO: GO:0006508; P:proteolysis; IEA.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Src_Fcpt.
 DR Pfam: PF00089; Trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYSPC; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 359 AA; 38943 MW; 4C14083C78233B37 CRC64;
 Alignment Scores:
 Pred. NO.: 6.83e-42 Length: 359
 Score: 891.00 Matches: 159
 Percent Similarity: 63.6% Conservative: 70
 Best Local Similarity: 44.2% Mismatches: 121
 Query Match: 26.7% Indels: 10

```

DB: 2 Gaps: 4
US-10-806-370-11 (1-1748) x Q4SPG0_TESTNG (1-359)

QY 614 CAGCGCTTGGCTGTGACGGGCTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGCTGC 673
DB 1 AenAlaThrHrHsCyAspGlyValArgAspCysThrLeuGlySerAspGluThrAlaCys 20
QY 674 GTGAGCTTTGACCTGGACAGAGCTCTGCTTAAATCTACTGAGTCTCCATCAAGTGG 733
DB 21 ValMetLeuGlyAsnAsp---AsnIleLeuGlnValIysThrSerGlnAspGlyArgPhe 39
QY 734 CTTCCTCATCTGTAGCAACAAGTGAATGACTCTTACTGAGAGAGACTTGCACCACTG 793
DB 40 LeuProValCysTyraAnGlyTTPAspGlySerLeuAlaGlySgluThrCysThrIleu 59
QY 794 GGTTCGAGAGTGTCAACCGAGCAACGAGGTTCCACAGGAGTTTGGCAACAGCTTC 853
DB 60 GlyPheArgAsnPhenylAlaThrAsnProSerThrSerGln-----ProLysSerSer 77
QY 854 TCATCTTGAAGATACAACTCCACCATCCAGAGAAAGCTCCAC-----AGG 898
DB 78 ProThrLeuThrIleAsnSerArgSerSerProTyrlLeuGlnGlyArgValAsnValSer 97
QY 899 TCTGAATGCCCTTCCACAGCGGTATATCTCCCTCAGGTTCCTCCAGTGGAGCTGAGGACC 958
DB 98 SerSerCysProGlyGlnGlnThrValAlaLeuGlnCysLeuAspCysGlyGlnArg 117
QY 959 ATGACCGGAGGATCGGAGGAGGCGGCTGCGCATGACAGTGGCGCTTGCAGAGT 1018
DB 118 SerThrSerArgIleIleGlyGlyAsnValAlaLysLeuGlnGlnThrProTyrGlnMet 137
QY 1019 AGTGTGACTTGGACCAACCATCTGTGAGAGGACGCTCATTTGACGCGCATGGAGT 1078
DB 138 ThrIleuThrPheArgGlySerHsValCysGlyGlyIleLeuIleSerProAspPheVal 157
QY 1079 CTGACGCGCGGACCTGCTTCTTGTGTAACCGGAGAGAGGCTCTG-----GAGGCGTGG 1132
DB 158 LeuThrAlaIleAlaHsCysPheProGlySerAsnIleLeuAlaIleGluAsnTyr 177
QY 1133 AAGGTGTACCGGCGACACAGAAACCTGCACAGTGTCTGAGGACGCTCCATTCGCGGAG 1192
DB 178 GluValIysIserGlyValGlnIserLeuAspLysLeuProLysProTyrLysValLysArg 197
QY 1193 ATCATCATCAACAGCAATTAACCGCATGAGAGGAGGACACTATGACATGCGCTCATCGG 1252
DB 198 IleLeuLeuSerGluLeuTyrAsnSerAspThrAsnAspTyrAspValAlaLeuLeuLys 217
QY 1253 CTGTCCAGCGCGCTGACCTGTCCGCTCACATCCACCTGCTGCTCCCGCATCATGGA 1312
DB 218 LeuAlaIleProValIlePheAspAspAsnValGlnProAlaCysLeuProSerArgAsp 237
QY 1313 CAGACCTTGGCCCTCATGAGAGACTGTGATCAAGCTTTGGCAGAGACGAGGAGACA 1372
DB 238 GlnIleLeuAlaProGlyThrGlnCysIstTyrThrIleGlyPheGlyThrThrGluAspGly 257
QY 1373 GATGACAGAGCATCTCCCTTCTCTCCGAGAGGAGTCACTCATCTGCACTTCAAGAAA 1432
DB 258 SerSerSerValIserLysSerIleuMetGluValSerValAsnIleIleSerAspThrVal 277
QY 1433 TGCATATCACTTGTGTCTATGACAGTTTACCTTACCCCAAGATGATGTGTCTGGGAC 1492
DB 278 CysAsnSerValThrValItyAsnIlyAlaValAlaThrLysAsnMetLeuCysAlaGlyAsp 297
QY 1493 CTTCGTGGGGGAGAGACTCTGTCCAGGAGAGACGCGGGGGGCTCTGTCTGTGAGACAG 1552
DB 298 LeuIysGlyGlyLysAspSerCysGlnGlyAspSerIlyGlyProLeuValCysGlnGln 317
QY 1553 AACACCGCTGTACTGTGAGAGTGTCAACAGCTGTGGGACAGGCTGTGGCAGAGAAAC 1612
DB 318 AspAspArgIstTyrValValGlyIleThrSerThrIleSerGlyCysGlyGlnAlaAsn 337
QY 1613 AAACCTGTGTGTACACCAAGTGAAGAGAGTTCTTCCCTGAGATTTAACAGCAAGATGAG 1672

```

```

DB 338 LysProGlyValTyrThrArgValSerSerValLeuProTyrIleTyrSerArgMetGln 357

RESULT 4
ID Q311V4_PANPA PRELIMINARY; PRT; 484 AA.
AC Q311V4;
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Transmembrane protease serine 2 (Fragment).
GN Name=TPRSS2;
OS Pan paniscus (Pygmy chimpanzee) (Bonobo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Pan.
NCBI_TaxID=9597;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Clark N.L., Swanson W.J.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License
CC -----
DR EMBL; DQ150500; AA282285.1; -, Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR002172; LDL_rcpt_A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Src_c-rcpt.
DR Pfam; PF00089; Tyrosin; 1.
DR SMART; SM00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLA; 1.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS50068; LDLRA_2; 1.
DR PROSITE; PS50287; SRCR_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Protease; Transmembrane.
FT NON_TER 1
FT TER 484
SQ SEQUENCE 484 AA; 52970 MW; CC7D4F9E1B8A0C32 CRC64;

Alignment Scores:
Pred. No.: 7.07e-36 Length: 484
Score: 784.50 Matches: 181
Percent Similarity: 51.1% Conservative: 72
Best Local Similarity: 36.6% Mismatches: 191
Query Match: 23.5% Indels: 51
DB: Gaps: 15

US-10-806-370-11 (1-1748) x Q311V4_PANPA (1-484)
QY 288 GGTCTTCATCCGCGAGGTGAT---CATCCGCGAGTCAAGCTCGTGAACAACCTCCCAA 344
DB 7 GlyProTyrTyrGluAsnHsIleGlyTyrGlnProGluAsnPro-----TyrProAlaGln 24
QY 345 CCAAGAGTGT-----ACCTTGTTAAGCAAACACCAAGTGGGGGCTGTACCATCTC----- 392
DB 25 ProThrValAlaProThrValTyrGluValIleProAlaGlnTyrTyrProSerProVal 44
QY 393 GATCATCTCTGCGCAGGTCAACAGCAACAGGCGGACCAAGGAGAGAGCCAGGTACGA 452
DB 45 ProGlnTyrAlaProArgValLeuThrGlnAlaSerAsnPro-----ValValArg 61

```



```

QY 453 GCCTGCCAAGTTGACCTGGGAGGCGCAGAACGCTACCGCTCATGGGTCGTC 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 MetcInProlyserProserGlyThrValCysThrSerLysThrLysAlaLeuLys 81
QY 513 -----TCCCTCTCAATGCCCTCG 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 LeuThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp 101
QY 531 TGGTTGGCTCATATCTCTTCCAGATTTCGGCAGGCGCACAGGATCAGGTACAG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 120
QY 591 AGCAGAGGAGAGAGCTGCCAAGACCGCTGCTCCGTGAGGAGGCTGTGACCTGCAAC 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 -----CysIleSerProSerAsnTrp---CysAspGlyValSerHisCysProG 136
QY 651 TGAAGAGTGAAGCTGGGCTGGCTGAGGCTTGACTGGGCAAGTCTCTGCTTAATCT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 LysGlyGluAspGluAsnArgCysValArgLeuTyrglyProAsnPheIleLeuGlnValT 156
QY 711 ACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGCAGACACTGGAATGACTCTACT 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 156 YrSerSerHisSerLysSerTrpHisProValCysGlnAspTrpAsnGluAsnTyrg 176
QY 771 CAGAGAAAGCTGCCAGCAGCTGGGTTCCGAGAGTCTCACCCGACACACAGGTTGCC 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 LysArgAlaAlaCysArgAspMetGlyTyrgLysAsnAsnPheTyrgSerSerGln---GlyI 195
QY 831 ACAGAGATTTTGCACACAGCTTCTCAATCTGAGATACACTCCACAC----- 877
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LevalAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 215
QY 878 --ATCCAGAAAGGCTCCACAGGTCTGAA--TGCCTTCCACAGCGATATCTCCCTCC 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 SpIeTyrgLysLeuLeuTyrgHisSerAspAlaCysSerSerLysAlaValAlaSerLeuA 235
QY 933 AGTGTCTCCACCTGGGCACTGAGG-----GCCATGACCGGGCGGATCTGTCGAGGG 983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 TgCysIleAlaCysGlyValAlaAsnLeuAsnSerSerArgGlnSerTrgIleValGlyG 255
QY 984 CGCTGGGCTCGGATGACAAAGTGGCTTGCAAGTGAAGTCACTCCGACACACACCA 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 255 LysSerAlaLeuProGlyAlaTrpTrpGlnValSerLeuHisValGlnAsnValHisV 275
QY 1044 TCTGTGAGGACGACCTCATGACGCCAGTGGTGTCTCACTGCCGCCACCTGCTTTCG 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 alCysGlyGlySerIleIleThrProGluTrpIleValThrAlaAlaHisCysVal---- 293
QY 1104 TGACCCGGGAGAGGTCTCGAGGGC-----TGAAGGTGACCGGGCACACAGCA 1154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 -----GluLysProLeuAsnAsnProTrpHisIleThrAlaPheAlaGlyIleLeuA 311
QY 1155 ACCTGACACCAAGTGGCTGAGGACGCC-----TCCATGTCGCGAGATCATCAACAGCA 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 TgGlnSerPheMetPheTyrgLysAlaGlyTyrgLysValGluLysValIleSerHisProA 331
QY 1209 ATTAACACCGATGAGAGGACGACTATGACATCGCCCTCATCGGCTGTCAAGCCCTGA 1268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 snTyrgAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 351
QY 1269 CCGCTGTCGCTCAATCCACCTGCTTCCCTCCCATGATGACAGACCTTTAGCTTCA 1328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetLeuLeuGluProG 371
QY 1329 ATGAGACCTGTGTGATCAGAGCTTTGGCAAGACAGGAGAGACATACAGACATCCC 1388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 LysGlnLeuCysTrpIleSerLysTrpGlyAlaThrGluGln---LysGlyLysThrSerG 390
QY 1389 CTTTCTCTCGGAGGTGAGGTCAATCTCATGACTTCAAGAAATGACATGACTTCTGG 1448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 LuValLeuAsnAlaAlaLysValLeuLeuIleGluThrGlnArgCysAsnSerArgTyrv 410

```

```

QY 1449 TCTATGACAGTACTTATACCCAGAGATGATGTGTCTGGGACCTTGTGGGCGAG 1508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 alTyrgAsnLeuIleIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 430
QY 1509 ACTCTGCCAGGAGAGACAGCGGGGGGCTCTGTGTGTGAGCAGACACACCGCTGTACC 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 spSerCysGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTrpTrpL 450
QY 1569 TGGCAGGTGTACACAGCTGGGGCGACAGCTGTGGCCAGAGAAACAAACCTGTGTACA 1628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTyrgProGlyValTyrg 470
QY 1629 CCAAGTACAGAAAGTCTTCCCTGAGTTTACAGCAAGATG 1669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 LysAsnValThrValPheThrAspTrpIleTyrgArgGlnMet 483

RESULT 5
Q311V3_9PRIM
ID Q311V3_9PRIM PRELIMINARY; PRT; 484 AA.
AC Q311V3_
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Transmembrane protease serine 2 (Fragment).
GN Name=TMPRSS2;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Clark N.L., Swanson W.J.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL, DQ150501; A282286.1; -, Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis; IEA.
DR InterPro; IPR002172; LDL_rcpt_A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_rcpt..
DR Pfam; PF00057; Ldl_recept_a; 1.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLA; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLR_1; 1.
DR PROSITE; PS00068; LDLR_2; 1.
DR PROSITE; PS50287; SRCR_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Protease; Transmembrane.
FT NON_TER 1
FT TER 484
SQ SEQUENCE 484 AA; 53153 MW; 3D197DA4442BB6C5 CRC64;

Alignment Scores:
Pred. No.: 8,05e-36 Length: 484
Score: 783.50 Matches: 184
Percent Similarity: 50.9% Conservative: 69
Best Local Similarity: 37.0% MisMatches: 189
Query Match: 23.5% Indels: 55
Gaps: 17

```

```

US-10-806-370-11 (1-1748) x Q3IIV3_gprim (1-484)
QY 288 GGTCCATATCCGGGAGGTCAT---CATCCGGCAGAGTCACTCGGTGACAACTCCCA 344
Db 7 GTPProtyrTygluamhnlsglytrglInProgluamPro-----TyrProalagln 24
QY 345 CCAGAGTGT-----ACCTGTGTAGAGCAACAGCAGTGGGGGCTGTACCCATCC----- 392
Db 25 ProthrValAlaProthrValAlaTygluValhlsProalaglnTytrTytrProserProval 44
QY 393 GATCATCTCCCTCCAGGTCAGACACAGACACAGGAGCCACAGGAGAGCCAGATGCA 452
Db 45 ProglIntyralaProargValleuthrGlInhlsaseramPro-----ValValArg 61
QY 453 GCCTGCCCAAGTTCACCTGGGAGGAGGAGCCAGAGAGCTTCCGCTCATCGGGGCGCG 512
Db 62 ThrGlnProlyserProserGlyThrValCysThrserlyThrLysLysAlaLeucyl 81
QY 513 -----TCCCTCCTCATTTGCCCTG 530
Db 82 lIethrleuthrleuglyThrPhleuValglValAlaLeuAlaAlaGlyLeuLeutr 101
QY 531 TGGTTTGGCTCATATCTCTTCCAGTTTGGCAGGGCCACACAGGGATCAGTACAGG 590
Db 102 LysPhemetglYserlySylserAspserGlyllegluCyAspserSerglyThr--- 120
QY 591 AGCAGAGGAGAGAGTGTCCCAAGCAGCTGTGCTGCTGAGCAGGGGTGGAGTCAAGC 650
Db 121 -----CyslleSerProserAmTrp---CysAspGlyValSerhlsCysProS 136
QY 651 TGAAGAGTGAAGAGTGGGCTGCGTGAAGTTTGAAGTGGAGCAAGTCTTGCTTAAATCT 710
Db 136 erglygluAspGluamhnlsglytrglInProamPhlleuGlInvalT 156
QY 711 ACTTGGGCTCTCCCATATCAGTGGCTTCCATCTGTAGACGACACTGGAAATCTCTACT 770
Db 156 ySersersglInArgLysSertrpHlsProValCysglInAspAspTrpAmhluamTyrg 176
QY 771 CAGAGAGAGCTGCGAGAGTTCGAGAGTTCGAGAGTTCACCGGACACCGAGTTGCC 830
Db 176 lyaArgAlaAlaCysArgAspmetglYtrglYAsamAmPhetyrSerSergln---GlyI 195
QY 831 ACAAGGATTTTGGCAACAGCTTCTCAATCTTGAGATCAACTCCACG----- 877
Db 195 leValAspAspserGlySerThrserPhemetlyLysLeuAmhlsSerAlaGlyamVala 215
QY 878 --ATCCAGAAAGCTCCACAGGTTGAA--TGCCTTCCACAGCGATATCTTCCCTCC 932
Db 215 spIetyrlyLysLysleuTytrhlsaserAlaCysSerSerglyAlaValValSerleuA 235
QY 933 AGTGTCCCACTGGGAGCTGAG-----GCCATGACCGCGGAGTGTGGAGGGG 983
Db 235 rGysylleAlaCysGlyValamhlsuamhlserserhlsrglnSerglyleAlaGlylv 255
QY 984 CGCTGGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTTCGACTTGGCAGCCACCA 1043
Db 255 alSerAlaLeuProglYAlaAlaTrpProtrpGlnValSerleuHlsValGlnamValhlsV 275
QY 1044 TCTGTGAGGAGCGCTCATTTGACGCCCAAGTGGGTCATCTGCGCGCCACTGCTTTCG 1103
Db 275 alCysgllyGlyserlleIethrProglIntyrlleValhlsAlaAlaHlsCysVal---- 293
QY 1104 TGACCCGGGAGAGAGTCTGAGAGG-----TGAAGGTGTAGCGGGGAGCAGCA 1154
Db 294 -----GluLysProleuamhlsamProtrpHlsStrpThrAlaPhemAlaGlylle---- 309
QY 1155 ACCTGACACAG-----TTGCTGAGGAGCGCC-----TCCATTTGCCAGATCATCATCA 1202
Db 310 --LeuArgGlnSerPhemetPhetyrGlualaglytyrGlnValGlnLysValhlsSerh 329
QY 1203 ACAGCATTTACACGAGTGAAGAGAGCACTATGACATCGCCCTCATGCGGCTGTTCACAG 1262
Db 1203 ACAGCATTTACACGAGTGAAGAGAGCACTATGACATCGCCCTCATGCGGCTGTTCACAG 1262

```

```

Db 329 lAsProamTyAspSerlyThrLysAsamhlsamhlsleAlaLeuMetlyLeuGlnlyuS 349
QY 1263 CCCTGACCCCTGTGCGGTCAACATCCACCTGTGCTTCCCTCCCATGATGAGACAGACTTTA 1322
Db 349 roLeuThrPhemAspLeuValLysProValCysLeuProamProglYmetMetleuG 369
QY 1323 GCCTCAATGAGACCTGTGTGATCAAGGCTTTGGCAAGACCGAGGAGACAGATGACAAGA 1382
Db 369 lufProglInleuCyetrpIleserGlytrpGlyAlaThrGlnGlu---LysglYser 388
QY 1383 CATCCCTCTCTCCCGGAGGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGACT 1442
Db 388 hrSerGluValleuamhlsAlaAlaArgValleuIleGluThrGlnLysCysamhlsEra 408
QY 1443 ACTTGCTATGACAGTACCTTACCCCAAGATGATGTGCTGGGAGCTTGCTGAGGAG 1502
Db 408 rglYrVallyrAspAmhlsleuIleThrProAlaMetIleCysAlaGlyPhleuGlnGly 428
QY 1503 GCAGAGACTCCTGCGCAGAGAGACAGCGGGGCGCTTGTCTGTGACAGACAAACGCT 1562
Db 428 enValAspserCysGlnGlyAspserGlyglYProleuValThrLeuLysamhlsValT 448
QY 1563 GGTACTGTGAGAGTGTACCAAGCTGGGACAGGCTGTGGCCAGAGAAACAACTGCTG 1622
Db 448 rprtleuIleGlyAspThrsertrpIleserGlyCysAlaLysAlaTyrgArgProglY 468
QY 1623 TGTACACCAAGTACAGAGATTTCTTCCCTGATTTTACAGACAGATG 1669
Db 468 alTyrglyamhlsValThrValPhemThrAspTrpIetyrArgGlnMet 483

RESULT 6
Q3IIV3_PANTR
ID Q3IIV3_PANTR PRELIMINARY; PRT; 484 AA.
AC Q3IIV3
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Transmembrane protease serine 2 (Fragment).
GN Name=TMPRSS2;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Pan.
OX NCBI_TaxID=9598;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Clark N.L., Swanson M.J.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
-----
EMBL, DQ150499; AA82284.1; -, Genomic DNA.
GO: GO:0016021; C: integral to membrane; IEA.
GO: GO:0016020; C: membrane; IEA.
GO: GO:0008233; F: peptidase activity; IEA.
GO: GO:0005044; F: scavenger receptor activity; IEA.
GO: GO:0004252; F: serine-type endopeptidase activity; IEA.
GO: GO:000508; P: proteolysis; IEA.
InterPro: IPR002172; LDL_rcpt_A.
InterPro: IPR001254; peptidase_S16.
InterPro: IPR001314; peptidase_S1A.
InterPro: IPR001190; Srcr rcpt_Pfam; PF00057; Ldl_recept_a; 1.
Pfam: PF00057; Ldl_recept_a; 1.
PRINTS: PR00722; CHMOTRYPsin.
SMART: SM00192; LDLa; 1.
SMART: SM00202; SR; 1.
SMART: SM00020; Tryp_Spc; 1.
PROSITE: PS01209; LDLRa_1; 1.
PROSITE: PS50068; LDLRa_2; 1.
PROSITE: PS50287; SRCR_2; 1.
PROSITE: PS50240; TRYPsin_DOM; 1.

```

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 484 AA; 53118 MW; 33AB1E471DE1344 CRC64;
 Alignment Scores:
 Pred. No.: 1,19e-35 Length: 484
 Score: 780.50 Matches: 180
 Percent Similarity: 50.9% Conservative: 72
 Best Local Similarity: 36.4% Mismatches: 192
 Query Match: 23.4% Indels: 51
 Gaps: 15
 DB: 2
 US-10-806-370-11 (1-1748) x Q311V5_PANTR (1-484)
 QY 288 GATCCATCCGCGAGGTCAT---CATCCGCGAGTCAGCTCCGTCGACAACTCCCAA 344
 DB 7 G1yProTyrTyrGluAenHISG1yTyrGlnProGluAenPro-----TyrProAlaGln 24
 QY 345 CCAGAGTGT-----ACCTGTAGACACACAGTGGGGGCTGACCCATCC----- 392
 DB 25 ProThrValAlaPProThrValTyrGluHisProAlaGlnTyrTyrProSerProVal 44
 QY 393 GATCATCTCTCCGAGGTCAGACACACAGACACAGGCGCACAGGAGAGAGCCAGTACGA 452
 DB 45 ProGlnTyrAlaPProArgValLeuThrGlnAlaSerAnPro-----ValValArg 61
 QY 453 GCGTCCCAAGTTCACCTGGCGGAGGCGCAGAGCAGCTACCGCTCATCGGTCGCGTC 512
 DB 62 MetGlnProLysSerProSerG1yThrValCysThrSerLysThrLysValLeuCys 81
 QY 513 -----TCCTCCTCATTCGCTCG 530
 DB 82 LeuThrLeuThrLeuGlyThrPheLeuValGlyAlaLeuAlaLeuAlaGlyLeuLeuTyr 101
 QY 531 TGGTTTGGCTCATATCTCTTCCAGTTCGTCGAGGCGCACACAGGATCAGTACAG 590
 DB 102 AsnPheMetGlySerLysCysSerAsnSerG1yLeuGlyCysAspSerSerG1yThr-- 120
 QY 591 AGCAGAGGAGAGAGCTGTCCCAAGCAGCTGTGCTGTGACGGGTGGTGGACAGC 650
 DB 121 -----CysIleSerProSerAsnTyr--CysAspGlyValSerHisCysProS 136
 QY 651 TGAAGAGTACGAGAGCTGGGCTGCTGAGGTTTGACTGGGACAACTCTGCTTAAATCT 710
 DB 136 erg1yGluAspGluAsnArgCysValAlaTyrGluTyrProAsnPheIleLeuGlnValT 156
 QY 711 ACTCTGGGCTCTCCCATCATGAGCTTCCCATCTGTAGCAGCACTGGATGATCTTACT 770
 DB 156 YrSerSerGlnArgLysSerTyrPheIleProValCysGlnAspAspTyrAsnGluAsnTyrG 176
 QY 771 CAGAGAGAGCTGTCCGAGAGCTGGGTTTCAGAGAGTGTCAACGGACAAACGAGTGGCC 830
 DB 176 LysArgAlaAlaCysArgAspMetG1yTyrLysAsnAsnPheTyrSerSerGln--GlyI 195
 QY 831 ACAGGAGATTTTGCACAGACTTCTCAATCTTGAGATCAACTCCACC----- 877
 DB 195 IeValAspAspSerG1ySerThrPheMetLysLeuAsnThrSerAlaGlyAsnValA 215
 QY 933 AGTGTCCCATCTGGGAGCTGAG-----GCCATGACCGGGCGGATCTGGGAGGGG 983
 DB 235 rGcysIleAlaCysGlyValAsnLeuAsnSerSerArgGlnSerArgIleValGlyLeu 255
 QY 984 CGTGGCTCCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTGGCAGCAGCACA 1043
 DB 255 IuSerAlaLeuProGlyValaTyrProTyrGlnValSerLeuHisValGlnAsnValHisV 275

QY 1044 TCTGTGAGGACAGCTCATTTAGACCCGAGTGGGNGTCACTGCCGCCACTGTTCTTG 1103
 DB 275 a1CysGlyGlySerIleIleThrProGluTyrIleValThrAlaAlaHisCysVal---- 293
 QY 1104 TGACCCGGAGAAAGCTCTGAGAGGC-----TGAAGGTGACCGGAGCAGCACA 1154
 DB 294 -----GluLysProLeuAsnAsnProTyrPheIleTyrPheAlaAlaGlyIleLeuA 311
 QY 1155 ACCTGCACCAAGTTCCTGAGAGCACC-----TCCATTCGCCGAGATCATCAACACA 1208
 DB 311 rGlnSerPheMetPheTyrGlyAlaGlyTyrGlnValGluLysValIleSerHisProA 331
 QY 1209 ATTAGACCGATGAGAGAGAGACACTAGATACATGCGCCCTCATGCGCCTCTCCAAAGCCCTGA 1268
 DB 331 enTyrAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 351
 QY 1269 CCTGTCCGCTCATCATCACACCTGCTGCTGCTCCCATGATGACAGACAGACCTTAACTCA 1328
 DB 351 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGluProG 371
 QY 1329 ATGAGACCTGCTGATCAGAGGCTTTGGCAAGACCAAGAGACAGATGACAAACATCC 1388
 DB 371 IuGlnLeuCysTyrPheIleSerG1yTyrGlyAlaThrGluGlu--LysGlyLysThrSerG 390
 QY 1389 CCTTCCCTCCGGAGAGCTGACAGGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG 1448
 DB 390 IuValIleuAsnAlaAlaLysValLeuLeuIleGluThrGlnArgCysAsnSerArgTyrV 410
 QY 1449 TCTATGACAGTTCATCTTACCCCAAGATGATGTGTGTGGGACCTTGTGGGGGAGAG 1508
 DB 410 a1TyrAspAsnLeuIleIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 430
 QY 1509 ACTCTCCCAAGGAGACAGCGGGGGCTCTTGTGTGTGAGCAGACAAACCGCTGTAC 1568
 DB 430 spSerCysGlnGlyAspSerG1yLysProLeuValThrSerLysAsnAsnIleTyrPl 450
 QY 1569 TGGAGAGTGCACAGAGTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTACA 1628
 DB 450 euIleGlyAspThrSerTyrGlySerGlyCysAlaLysValTyrAspGlyValTyrG 470
 QY 1629 CCAAGTGTACAGAAATCTTCCCTGATTTTACAGCAAGATG 1669
 DB 470 LysAsnValThrValPheThrAspTyrIleTyrArgGlnMet 483
 RESULT 7
 ID TMS2 MOUSE STANDARD; PRT; 490 AA.
 AC Q9UIOB: Q9UIK4: Q9OY82;
 DT 11-JAN-2001, integrated into UniProtKB/Swiss-Prot.
 DT 11-JAN-2001, sequence version 2.
 DT 07-MAR-2006, entry version 52.
 DE Transmembrane protease, serine 2 (EC 3.4.21.-) (Epitheliasin) (Plasmic
 DE transmembrane protein X) (Contains: Transmembrane protease, serine 2
 DE non-catalytic chain; Transmembrane protease, serine 2 catalytic
 DE chain).
 GN Name=Tmprs2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [mRNA], AND TISSUE SPECIFICITY.
 RC STRAIN=BALB/c;
 RX MEDLINE=21104370; PubMed=11169526;
 RX DOI=10.1002/1096-9896(2000)9999<:AID-PATH743>3.0.CO;2-T;
 RA Vaarala M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Vahko P.T.,
 RT "Expression of transmembrane serine protease Tmprs2 in mouse and
 human tissues".
 RL J. Pathol. 193:134-140(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [mRNA].
 RA Han J., Kim S.;

RT "Putative transmembrane protease X.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE (mRNA).
 RC STRAIN-BALB/c;
 RX PubMed=10683448; DOI=10.1016/S0014-5793(00)01196-0;
 RA Jacquinet E.J., Rao N.V., Rao G.V., Hoidal J.R.;
 RT "Cloning, genomic organization, chromosomal assignment and expression
 of a novel mosaic serine proteinase: epithelisin.";
 RL FEBS Lett. 468:93-100(2000).
 RN [4]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA).
 RC STRAIN=129; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshlyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguailano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pailey J., Helton E., Kesteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
 CC cleavage and secreted (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed mainly in prostate and kidney.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family.
 CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -1- SIMILARITY: Contains 1 SRCR domain.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License

KW Signal-anchor; Transmembrane; Zymogen.
 FT CHAIN 1 253 Transmembrane protease, serine 2 non-
 FT catalytic chain.
 FT FT /Ftrid-PRO_0000027857.
 FT CHAIN 254 490 Transmembrane protease, serine 2
 FT catalytic chain.
 FT FT /Ftrid-PRO_0000027858.
 FT FT Cytoplasmic (Potential).
 FT FT Signal-anchor for type II membrane
 FT FT protein (potential).
 FT FT Extracellular (potential).
 FT FT LDL-receptor class A.
 FT FT SRCR.
 FT FT Peptidase S1.
 FT FT Charge relay system (By similarity).
 FT FT Charge relay system (By similarity).
 FT FT Charge relay system (By similarity).
 FT FT ACT_SITE 439 433 Cleavage (potential).
 FT FT ACT_SITE 439 433 Cleavage (potential).
 FT FT SITE 253 254 N-linked (GlcNAc...) (Potential).
 FT FT CARBOHYD 111 111 N-linked (GlcNAc...) (Potential).
 FT FT CARBOHYD 212 212 N-linked (GlcNAc...) (Potential).
 FT FT CARBOHYD 474 474 N-linked (GlcNAc...) (Potential).
 FT FT DISULFID 76 125 By similarity.
 FT FT DISULFID 119 138 By similarity.
 FT FT DISULFID 132 147 By similarity.
 FT FT DISULFID 171 230 By similarity.
 FT FT DISULFID 184 240 Interchain (between non-catalytic and
 FT FT DISULFID 243 363 catalytic chains) (By similarity).
 FT FT DISULFID 279 295 By similarity.
 FT FT DISULFID 408 424 By similarity.
 FT FT DISULFID 435 463 By similarity.
 FT FT DISULFID 122 122 S -> L (in Ref. 3).
 FT FT CONFLICT 178 178 S -> H (in Ref. 3).
 FT FT CONFLICT 320 320 Y -> H (in Ref. 1).
 FT FT CONFLICT 474 474 N -> D (in Ref. 1).
 FT FT SEQUENCE 490 AA; 53480 MW; 07D2B03E4DBA1A9 CRC64;
 SO
 Alignment Scores:
 Pred. No.: 1,76e-35 Length: 490
 Score: 777.50 Matches: 180
 Percent Similarity: 47.4% Conservative: 81
 Best Local Similarity: 32.7% Mismatches: 201
 Query Match: 23.3% Indels: 89
 DB: 1 Gaps: 15
 US-10-806-370-11 (1-1748) x TMPS2_MOUSE (1-490)
 QY 89 GCTGGGACACCTCCAGCGCGGATCTCCAGCCGAGCATCTCCAGCCGAGCATCTCCA 148
 Db :::::|||||
 QY 5 SerGlySerProProGly---IleGlyProCysrTygIuLenhIsgIyTygIenSerGlu 23
 Db :::::|||||
 QY 149 GCTGGGACACCTCCAGCGCGGATCTCCAGCCGAGCATCTCCAGCTGGTACACTCCA 208
 Db :::::|||||
 QY 24 HisIleCysProPro-----ArgProPro 31
 Db :::::|||||
 QY 209 GCGCGGACATCTCCAGCGCGGATCTCCAGCCGAGCATCTCCAGCCGAGCATCTCCG 268
 Db :::::|||||
 QY 32 ValAlaProLenhIyTyraenLeuTyProAlaGlnTyTyPro----- 46
 Db :::::|||||
 QY 269 GCTTGGCATCATCTTCCAGAGTCTCATCCGGCAGGTCATCATCCGACAGTCACTCG 328
 Db :::::|||||
 QY 46 ----- 46
 Db :::::|||||
 QY 329 GTGACACCTCCCAACCAAGTGTACTTGTAGACCAACCAAGTGGGGGTGATACC 388
 Db :::::|||||
 QY 47 -----SerProValProGlnTyraLeProhrgIlehnTrginalaSerThSer 63
 Db :::::|||||
 QY 389 ---ATCCGATCATCTCTCCGACAGTCA---GCACCAACCAACCAAGCCACCAAGAGAGC 442
 Db :::::|||||
 QY 64 ValIleHisThrHisProLysSerSerGlyAlaProCysThrSerLysSerLysSer 83
 Db :::::|||||
 QY 443 CCAGGTACGAGCTGCGCCCAAGTTCACCTGGCGGAGGAGCCAGAAAGACTACCGCTCATC 502
 Db :::::|||||


```

QY 288 GGTCCATCCGAGGATCAT---CATCCCGCAGGTGAGCTCCGATGCAACCTCCCA 344
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 7 G1yProG1yTg1uAenH1eG1yTg1nProG1uAenP-----TyProlAg1n 24
QY 345 CCAAGAGT-----ACCTTTAGAGCAACAGAGGGGGCTGTACCCATCCGATCAT 398
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 25 ProHrValAlaProAenValTg1uValnH1eProAlaG1nTyTyrProSer----- 42
QY 399 CTCCTGCAGGTGACAGCAGCAGAGGGCCACAGGGAGAGAGCCAGTA----- 449
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 43 -----ProValProlG1nTyTyrHrProAenValLeuThrHis 54
QY 450 -----CAGCCTGC-----CAAAGTACCTGGCGAGGAGCCAGAGCAGC 491
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 55 AlaSerAenProAlaValAlaCyAenG1nProLysSerProSerG1yThrValCyAenThSer 74
QY 492 TACGGTCATCGGGTGGCTGC----- 512
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 75 LyAThrLyValAlaLeuCyAenValThrMetThrLeuG1yAlaValLeuValG1yAlaAla 94
QY 513 ---TCCTCTCATTCCTCCCTGGTGGTTCGTCATCATCTCTCCAGTTCCGAGAGGCC 569
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 95 LeuAlaAlaG1yLeuLeuThrLyPheMetG1ySerLyCyAenAerSerG1y1eG1u 114
QY 570 ACACAGGATCAGTACAGAGCAGAGAGAGAGCTCCCAAGCAGCCTGTCGCTGTG 629
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 115 CyAaerSerSerG1yThr-----Cys1eSerSerSerAenTrp--CysA 129
QY 630 ACGGGGTGTGATGTCAGAGCTGAAGATGACAGAGCTGGCTGGCTGAGTTTACTGG 689
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 129 ePrg1yValSerH1eCyAerProAenG1yAlaAerG1yAenA1gCyAenAlaG1yLeuTyP 149
QY 690 ACAAGTCTCTCTTAATTAATTAATCTCTGGCTCCCAATCATGTGCTCCATCTGTAGCA 749
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 149 roAenPhe1eLeuG1nValTyTyrSerG1nA1gLyAerTyTyrH1eProAlaCyAenA1g 169
QY 750 GCAACTGGATGACTCTCTCAAGAAAGACCTGCGACAGCTGGCTTCCAGAGCTGTC 809
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 169 ePRAerTPRAenG1nA1nTyTyrAlaG1nAlaCyAenA1gAerMetG1yTyTyrA1nSerP 189
QY 810 ACGGAGACACCGAG--GTTCGCCAGAGGATTTTGGCAAGCTTCGATCTTGAGAT 866
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 189 heTyTyrSerG1nG1y1eValAerAenSerG1yAlaThrSerPheMetLyAenA1nT 209
QY 867 AC-----AATCCACATCCAGAAAGACCTCCACAGCTGAA---TGCCTTCCC 914
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 209 hrSerAlaG1yAenValAerP1eTyTyrLyAerLyAerTyTyrH1eSerA1aCyAenSerE1 229
QY 915 AGCGGTATCTCTCCCTCAAGTGTCCACCTCCGAGCTGAGG-----GCCATGACCG 965
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 229 yValA1aValSer1eLeuA1gCyAenValA1gSerA1nLeuSerA1g1nS 249
QY 966 GCGGATCGTGGAGGGCGCTGCGCTCCATGACAGTGGCTTGGCAATGAGTGGC 1025
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 249 e1aG1yG1eValG1yG1nAenH1aLeuLeuG1yAla1TrpTyrProlnA1e1eH1eU 269
QY 1026 ACTTCGACACACACATCTGTGAGAGCAGCTCATTTAGCAGCCAGTGGCTGCTCACTG 1085
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 269 1eValG1nAenH1a1eValCyAenG1ySer1e1e1eThrProln1Trp1eValH1a 289
QY 1086 CCGGCACTGCTTTGTGTGACACCGGAGAAAGTCTCGAGAGGC-----TGAAAGG 1136
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 289 1aAlaH1eCyAenVal-----G1uLyAerProLeuAenSerPro1Trp1n1TrpH1a 305
QY 1137 TGTAAGCGGGAGACAGACCTGACACAGCTGG-----CTGAGAGCAGCTCCATTCGGC 1190
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 305 1aPheValG1yThrLeuA1g1nSerSerMetP1eTyTyrG1uLyAerG1nH1eA1gValG1uL 325
QY 1191 AGATCATCATCAAGACATTAACCCAGTACAGAGAGAGACATATGACATCCGCTCATGC 1250
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 325 yVal1e1eSerH1eProAenTyTyrAerSerLyAerTyTyrA1nA1nAerP1e1a1eLeuMetL 345
QY 1251 GAGTGTCAAGACCCCTGACCTGTCCGCTCCATGACACCTGCTTGGCTCCCATGATG 1310

```

```

Db 345 yLeuH1eThrProLeuThrPheAenG1yVal1yPheProValCyAenProAenP1rog 365
QY 1211 GACAGACCTTTAGCCTCAATGAGACCTCGGTGATCAACAGCTTTGGACAGAGGGAGA 1370
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 365 1yMetMetLeuG1nProG1uG1nH1eCyAerP1eSerG1yTg1yAlaThrG1nG1u 384
QY 1371 CAGATGACAAAGACATCCCTCTCCCGAGGTGACAGTCAATCTCATGACTTCAAGA 1430
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 385 --LyG1yLyAerThSerAerP1eValLeuAenH1a1eMetMetP1eLeu1eG1uProA1rG 404
QY 1431 AATGCAATGACTTACTTGTCTATGACAGTACCTTACCTTACCCCAAGATGATGTCTGGG 1490
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 404 rgCyAenAenH1yTyTyrValTyAerP1yLeu1eThrProAlaMet1eCyAenA1gP 424
QY 1491 ACCTTCGTTGGGCGAGAGACTCTCCGAGGAGAGACAGCGGGGGCTCTGTGTGAGC 1550
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 424 heLeuG1nG1yThrValA1aSerP1eCyG1nG1yAerSerG1yG1yProLeuValThrLeu 444
QY 1551 AGAACCAACCGCTGTACTCTGACAGTGTCAACAGCTGGGCGACAGGCTGTGGCAGAGA 1610
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 444 yAenAerP1eVal1TrpTyP1eU1eG1yAerP1eThSerTrpG1ySerG1yCyAenA1nA1a 464
QY 1611 ACAAACTGGTGTGTATACCAAAAGTGCAGAGAATCTTCCCTGGATTTTACGACAAGTG 1669
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 464 eN1aG1yValTyG1yA1nValThrValPheThrAerP1eTyTyrA1rG1nMet 483

RESULT 9
ID Q6P7D7_RAT PRELIMINARY; PRT; 490 AA.
AC Q6P7D7/
DT 05-JUL-2004, integrated into UniProtKB/TREMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 18.
DE Transmembrane protease, serine 2.
GN Name=Trpser2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP NCBIOTIDE SEQUENCE.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Canninci P., Mullaly S.J.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnaracne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huilyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kesteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez M.T., Krzywicki M.I., Skalska U., Smallue D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NCBIOTIDE SEQUENCE.
RC TISSUE=Prostate;
RA Strauberg R.;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License

```

| | | |
|--|----------|------------------|
| Alignment Scores: | | |
| pred. No.: | 3,156-35 | length: 490 |
| Score: | 773.00 | Matches: 180 |
| Percent Similarity: | 49.5% | Conservative: 78 |
| Best Local Similarity: | 38.5% | Mismatches: 204 |
| Query Matchn: | 23.2% | Indels: 39 |
| DB: | 2 | Gaps: 15 |
| DS-10-806-370-11 (1-1748) x Q6P7D7 RAT (1-490) | | |

```

Db      139  oasnGlyGlnaspIuhsnRnrgCysValArgLeuTyGlyThnSerPheThrLeuGlnVala 159
Qy      709  CTACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGACGACACTGGAAATGACTCTCTA 768
Db      159  lTySerSerGlnArgLysalatrPyrProvalCysGlnAspSerPhePngInguSerTy 179
Qy      769  CTCAGAGAAACCTGGACGAGCTGGGGTTTCGAGAGGCTCAACCGGCAACCGAG---GT 825
Db      179  rGlyArgAlaAlaCysLysAspMetGlyTyLysAsnSerPheTySerSerGlnGlyI 199
Qy      826  TGCCCAAGAGGATTTTCCCAACAGCTTCTCAATCTTAGATAC-----AACTCCAC 876
Db      199  eProAspGlnSerGlyAlaThrSerPheMetLysLeuAsnValSerAlaGlyAsnVala 219
Qy      877  CATCCAGGAAAGCTCCACAGGTTGAA---TGCCCTTCCACAGCGGTAATCTCCCTCCA 933
Db      219  pLeuTyLysLysLeuTyrhIHisSerAspSerCysSerSerArgMetValValSerLeuAr 239
Qy      934  GTGTTCCCACTGGCGGACTAGAGGGCCAG-----ACCGGGCGGATTCGTGGAGGGGGCT 987
Db      239  gCysIAsglucysGlyValArgSerValArgArgGlnSerArgIleValGlyGlySerTh 259
Qy      988  GGCTCGGATPAGCAAGTGGCCCTTGGCAAGTAGCTGCACCTTGGGACCAACCCATCTG 1047
Db      259  rAlaSerProGlyAspTrpProTrpGlnValSerLeuHiValGlnGlyIleHisValCy 279
Qy      1048  TGAAGGACAGCTCATTTGACGCCAGTGGGTGCTCAGTCCGCGCCACTGCTTCTTGTCAG 1107
Db      279  sGlyGlySerIleIleThrProGluTrpIleValThrAlaAlaHisCysVal----- 296
Qy      1108  CCGGAGAGAGTCTCGAGGGC-----TGAAAGTGTACCCGGGGAC----- 1150
Db      297  ---GluGlnProLeuSerSerProArgTyTTrpThrAlaPheAlaGlyIleLeuLysG 315
Qy      1151  -----AGCAACCTGCACAGCTTCCGTAGGACACCTCCATGCCAGAT 1194
Db      315  nSerLeuMetPheTyrgLysSerAlaGlnSgl-----ValGluLysVa 329
Qy      1195  CATCATCAACAGCAATTTACACCCGATGAGAGGACACTATGACATCGCCCTCATGCGCT 1254
Db      329  lIleSerHisProAsnTyhAspSerLysrThnLysAsnAsnAspIleAlaLeuMetLys 349
Qy      1255  GTCCAAAGCCCTGACCCCTGTCGCCCTGCATATCCACCTGCTTGCTCCCATGATGACGA 1314
Db      349  uGlnThrProLeuAlaPheAsnAspValValLysProValCysLeuProAsnProGlyMe 369
Qy      1315  GACCTTTAGCTCATATAGACCTGCTGCATGACAGCGTTGGCAAGCAACGAGAGACGA 1374
Db      369  tMetLeuAspLeuAlaGlnGlyStrpIleSerGlyTyrPcLysAlaThrTyrgLnu--- 388
Qy      1375  TGACAAGACATCCCTCTCTCCGAGAGGAGTGCAGAGTCAATCTCATGCACTTCAGAAATG 1434
Db      388  sGlyLysThrSerAspValLeuAsnAlaAlaMetValProLeuIleGluProSerLysCy 408
Qy      1435  CAATGACTACTTGGCTATAGACAGTTACTTACCCCAAGATGATGTGTGCTGGGCACT 1494
Db      408  sAsnSerLysTyrlleTyhAsnAsnLeuIleThrProAlaMetIleCysAlaGlyPhele 428
Qy      1495  TCGTGGGGGCGAGAACTCTGCCGAGGAGACAGCGGGGGGCTTTGTCTGTACGACGAA 1554
Db      428  uGlnGlySerValAspSerCysGlnGlyAspSerGlyProLeuValThrLeuLysAs 448
Qy      1555  CAACCGCTGGTACCTGGCAGAGGTGCACACAGTGGGCGACAGGCTGTGGCCAGAAACAA 1614
Db      448  nGluIleTrpTrpLeuIleGlyAspThrSerTrpGlySerGlyAlaLysAlaTyhAr 468
Qy      1615  ACCTGTGTGTATACCAAAAGTACAGAAAGTTCTTCCCTGATTTTACAGACAGATGAGAG 1674
Db      468  gProGlyValTyrgLysAsnValThrValPheThrAspTrpIleTyrgInguInmetArgAl 488
Qy      1675  C 1675
Db      488  a 488

```


RESULT 10
Q3UKES3 MOUSE PRELIMINARY: PRT: 490 AA.
ID Q3UKES3
AC Q3UKES3
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 21-FEB-2006, entry version 7.
DE 14 days pregnant adult female placenta cDNA, RIKEN full-length
DE enriched library, clone:1510026G04 product:transmembrane protease,
DE serine-2, full insert sequence.
GN Name=Trps2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090.
RN [1]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Methods Enzymol. 303:19-44(1999).
RU [2]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX PubMed=1641072; DOI=10.1126/science.1112014;
RA Carninci P., Katayama T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Balci V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilmink L.G., Aldini J.E., Allen J.E.,
RA Amesl-Imphibato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Barsal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chaik A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Della E., Delympie B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engelstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Fukui S., Gariboldi M.,
RA Georghiou-Hemmings P., Gingras T.R., Gojobori T., Green R.B.,
RA Guetlich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnicki L., Iacono M., Ikeo K., Iwama A., Iehikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kikano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottacini-Taber S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J.F., Reid J.F., Ring B.Z., Ringwald M.,
RA Rest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schombach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
RA Tamajo A.K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yeig K.,
RA Yamashita H., Zabarovsky E., Zharov R., Zimmer A., Hide W., Bull C.,
RA Grimmond S.M., Teasdale R.D., Liu B.T., Brusic V., Quackenbush J.,
RA Mhlesetec C., Mettlick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imanura K., Itoh M., Kato T., Kawai H., Kawagashita N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimiya N.,
RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome";
RL Science 309:1559-1563(2005).
RN [3]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX PubMed=1641073; DOI=10.1126/science.1112009;
RG (Genome Network Core team) and the FANTOM Consortium;

RT "Antisense Transcription in the Mammalian Transcriptome";
RL Science 309:1564-1566(2005).
RN [4]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=22354683; PubMed=12468851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nishio T., Oseko N., Saito R., Suzuki H., Yamashita H., Kiyosawa H.,
RA Yeig K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,
RA Bardelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,
RA Schirral L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
RA Grimmond S., Guetlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kodzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,
RA Verardo R., Wagner L., Wallesed C., Wang Y., Watanabe Y., Wells C.,
RA Wilmink L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazune N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yamashita A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Binnery E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RU [5]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamashita H.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Flieschmann W., Gaasterland T., Giesi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nishio T., Oseko N., Quackenbush J.,
RA Schirral L.M., Strubli F., Suzuki R., Tomita M., Wagner L., Watanabe T.,
RA Sakai K., Okido T., Furuno M., Konno H., Bardelli R., Barin G.,
RA Blake J., Botelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momparati P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schombach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [6]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP Nucleotide sequence.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

QY 1400 GAGGTGACGTCATTCGACTTCAGAAATGCAATGACTTGTCTATGACAGT 1459
 Dp 397 AAlaMetValProLeuIleGluProSerIleCysAseSerIleTyIleTyrAseAse 416
 QY 1460 TACCTTACCCCAAGATGATGTGTCTGGGACCTTGTGGGGCAGAGATCTCTCCAG 1519
 Dp 417 LeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlySerValAseSerCysGln 436
 QY 1520 GGAGACAGCGGGGGCCCTTGTCTGTGTGAGCAACAACCCGTGATCTGGCAGGTGTC 1579
 Dp 437 GlyAseSerGlyGlyProLeuValThrLeuYsAseGlnIleTyrPheIleGlyAse 456
 QY 1580 ACCAGCTGGGCGACAGCTGTGGCCAGAAACAACCTGTGTGTATACCAAGTACA 1639
 Dp 457 ThrSerTrpGlySerGlyCysAlaIleYsAlaLeuArgProGlyValTyrGlyAseValThr 476
 QY 1640 GAAGTCTTCCCTGATTTACAGCAAGTGGAGAGC 1675
 Dp 477 ValPheThrAspTrpIleTyrIleGlyGlnMetArgAla 488

RESULT 11
 Q7TN04_MOUSE PRELIMINARY, PRT, 490 AA.
 AC Q7TN04_MOUSE integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2003, sequence version 1.
 DT 21-FEB-2006, entry version 16.
 DE Transmembrane protease, serine 2.
 GN Name: Tmpres2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/Fl;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA DiCicco L., Marzella K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapellato M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Ueding T.B., Toshiyuki S., Cavinini P., Prange C.,
 RA Raba S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman B., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallus D.B.,
 RA Schenck A., Schein J.B., Jones S.J.M., Maira M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/Fl;
 RX Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 DR HSPB, BC054348; AAHS4348.1; -; mRNA.
 DR HSPB, P00760; IEXX.
 DR Ensembl; ENSMUSG0000000385; Mus musculus.
 DR MGI, MGI:1354381; Tmpres2.
 DR GO, GO:0016021; C:integral to membrane; RCA.

DR InterPro; IPR002172; LDL_rcpt_A.
 DR InterPro; IPR001254; Peptidase_S1_S6.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001190; Src_rcpt.
 DR Pfam; PF00057; Ldl_recept_a; 1.
 DR Pfam; PF00530; SRCR; 1.
 DR Pfam; PF00089; Tyrosin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00192; LDLA; 1.
 DR SMART; SM00202; SR; 1.
 DR SMART; SM00020; Tryp_spc; 1.
 DR PROSITE; PS01209; LDLR_1; 1.
 DR PROSITE; PS0066; LDLR_2; 1.
 DR PROSITE; PS50287; SRCR_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydroxylase; Membrane; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 490 AA; 5352 MW; 54650B02841765A CRC64;

Alignment Scores:
 Pred. No.: 4,35e-35 Length: 490
 Score: 770.50 Matches: 174
 Percent Similarity: 47.3% Conservative: 87
 Best Local Similarity: 31.5% Mismatches: 200
 Query Match: 23.1% Indels: 91
 DB: 2 Gaps: 13

US-10-806-370-11 (1-1748) x Q7TN04_MOUSE (1-490)

QY 89 GCTGGAGACCTCCAGGCGCGGATCTCCAGCCAGCATCTCCAGCCAGCATCTCCA 148
 Dp 5 SerGlySerProProGly---11eGlyProCysTyrGlnAseIleGlyTyrGlnSerGlu 23
 QY 149 GCTGGAGACCTCCAGGCGCGGATCTCCAGCCAGCATCTCCAGCCAGCATCTCCA 208
 Dp 24 HisIleCysProPro-----ArgProPro 31
 QY 209 GAGCGGAGCATCTCCAGGCGCGGATCTCCAGCCAGCATCTCCAGC-----CGG 259
 Dp 32 ValAlaProAseGlnGlyTyrAseLeuTyrProAlaGlnTyrTrpProSerProValProGln 51
 QY 260 GCATCTCCGCGCTGGCATACCTTCAGAGCTCTCATCCGAGGATCATATCCGCCAGG 319
 Dp 52 TyrAlaProArgIleThrThrGlnAlaSerThrSerValIleThrThrProLysSer 71
 QY 320 TCGAGCCGCGTACCACTCCCAACCAAGAGTGTACTGTGTAGAGCAACAGAGTGGG 379
 Dp 72 SerGlyAlaLeuCysThrSerLysSerLysSerLysSerLysSerLysLeuAlaLeuGly 91
 QY 380 GCTGTACCCATCCGATCATCTCTGCCAGGTACAGCACCAACCAAGGCGCACCGAGAG 439
 Dp 92 ThrVal----- 93
 QY 440 AGCCCAAGTACAGAGCTGCCCAAGTTCACTGCGGAGGAGGAGGAGGAGGAGTACCCGCTC 499
 Dp 93 ----- 93
 QY 500 ATGGGAGGAGTCTCTCTCATTTGCCCTGATGTTTGGTCATCATCTTCCAGTTCC 559
 Dp 94 -----LeuThrGlyAlaAlaValAlaValAlaValLeuLeuTrpArgPhe 107
 QY 560 TGGCAGGCGCACAGAGATCAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 613
 Dp 108 TrpAseSerAseCysSerThrSerGlnMetGlnCysGlySerGlnTyrGlnIleSer 127
 QY 614 CAGCGCTGTGCTGTGACGCGGAGTGTGAGTGTGACGAGTGTGAAGAGTGTGAGGCTGC 673
 Dp 128 SerSerLeuTrpCysAseArgGlyValAlaIleCysSerProAseGlnGlyAseArgCys 147
 QY 674 GTGAGTTTACGTGGAGCAAGTCTCTGTTAAATGATCTAGTGTGGCTCTCCATCAATGG 733
 Dp 148 ValArgLeuTyrGlnGlnSerPheIleLeuGlnValTyrSerSerGlnArgValArgTyr 167

```

QY 734 CTTCCCATCTGTAGCAGCACTGGAAATGACTCTTACTCAGAGAAAGCCTGCGCAGAGCTG 793
DB 168 TyrProValCysGlnAspAspTrpSerGlnSerGlnValAlaCysAlaAspMet 187
QY 794 GATTTCGAGAGTGTGCTACCGGCAACCGAG--GTTGCCCAAGGATTTTGGCAACAGC 850
DB 188 GlnTyrTrpAlaMetPheTyrSerSerGlnGlnTyrProAspGlnSerGlnAlaThrSer 207
QY 851 TTCTCAATCTTGAGATACACTCC-----ACCATCCAGAAAGCCTCCACAGC 898
DB 208 PheMetLysLeuAlaMetValSerSerGlnValAlaSerLeuTyrLysLeuTyrHisSer 227
QY 899 TCTGAATGCTCTCCAGCGGATATATCTCCCTCCAGTTTCCCACTGGGAGCTAGCGGCC 958
DB 228 AspSerCysSerSerArgMetValValSerLeuAlaGlyGlnGlyValAlaArgSer 247
QY 959 ATG-----ACCGGCGGATGATGAGGAGGCGCTGCGGATGAGCAATGAGCTTGG 1012
DB 248 ValLysArgGlnSerArgLysValGlyLysLeuAlaSerProGlyLysTrpProTrp 267
QY 1013 CAACTGAGTCTGCACTTGGGACCAACCCACATCTGTGGAGGACGCTCATTGACGCCAG 1072
DB 268 GlnValSerLeuHisValGlnGlyValHisValCysGlyGlySerLysLeuThrProGln 287
QY 1073 TGGGTCTCTACTGCGGCCCACTGCTTTCTTCTGACCCGGAGAGAGTCTTGAGAGGC-- 1129
DB 288 TrpLeuValThrAlaAlaHisCysVal-----GlnGlnProLeuSerSerPro 303
QY 1130 -----TGAAGAGGTATACGGCGGACCC-----AGCAACTGTG 1159
DB 304 ArgTyrTrpThrAlaPheAlaGlyLysLeuArgGlnSerLeuMetPheTyrGlySerArg 323
QY 1160 CACCAATGCTGCTGAGGACGCTCCATTGCGAGATCATCATCAACAGCAATTAACCCGAT 1219
DB 324 HisGln-----ValGlnLysValLysSerHisProAlaThrAspSer 337
QY 1220 GAGAGAGCACTATGACATGCGCTCATGCGGCTGTCGAGCCCTGACCTGCGCT 1279
DB 338 LysThrLysAsnAspAlaLeuMetLysLeuGlnThrProLeuAlaPheAsnAsp 357
QY 1280 CACATCCACCTGCTTGGCTCCCATGATGATGAGCAAGCTTATGCTCATGAGAGCTGC 1339
DB 358 LeuValLysProValCysLeuProAsnProGlyMetMetLeuAspLeuAspGlnLys 377
QY 1340 TGGATCACAGGCTTTGGCAAGACCAAGAGCAGATGACAGACATCCCTTCCCTCGG 1399
DB 378 TrpLysSerGlyTrpGlyAlaThrTyrGln---LysGlyLysThrSerAspValLeuAsn 396
QY 1400 GAGGTGACGTCATCTCATGCTTCAAGAAATGCAATGACTACTTGGCTCATGACAGT 1459
DB 397 AlaAlaMetValProLeuLysLeuProSerLysCysAsnSerLysTyrLysAsnAsn 416
QY 1460 TACCTTACCCCAAGATGATGATGCTGGGAGCTTGGGAGGAGAGAGCTCCGCGCAG 1519
DB 417 LeuLysThrProAlaMetLysCysAlaGlyPheLeuGlnGlySerValAspSerGln 436
QY 1520 GGAGACAGCGGGGCGCTCTTGTCTGTGAGAGCAACCGCTGCTACTGCGCAGTGTGTC 1579
DB 437 GlyAspSerArgLysGlyProLeuValThrLeuLysAsnGlyLysTrpLeuLysGln 456
QY 1580 ACCAGCTGGGAGCAGGCTGTGGCCAGAAACAACTGGTGTGTACCAAGTGTACA 1639
DB 457 ThrSerTrpGlySerGlyCysAlaLysAlaLeuAspProGlyValTyrGlyAsnValThr 476
QY 1640 GAAGTCTTCTCGATTATACAGAGATGAGAGC 1675
DB 477 ValPheThrAspTrpLysGlnGlnMetAlaGala 488

```

RESULT 12
O6GT7 HUMAN
ID O6GT7 HUMAN PRELIMINARY; PRT; 492 AA.
AC O6GT7

DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Transmembrane protease, serine 2 variant.
GN Name=TPRSS2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Krausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.; to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA MEDLINE=94117032; PubMed=8125298; DOI=10.1016/S0378-1119(97)00411-3;
RX MEDLINE=94117032; PubMed=8125298; DOI=10.1016/S0378-1119(94)90802-8;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligonucleotides.";
RL Gene 138:171-174(1994).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=96038986; PubMed=9373149; DOI=10.1016/S0378-1119(97)00411-3;
RA Suzuki Y., Yoshitomo K., Maruyama K., Suyama A., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RT end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
RA Tanaka A., Yokoyama S.;
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC
DR EMBL; BC051839; AAHS1839.1; -; mRNA.
DR EMBL; AK222784; BAD9504.1; -; mRNA.
DR Ensembl; ENSG00000184012; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.

DR GQ: 0006508; P: proteolysis; IEA.
 DR InterPro: IPR002172; LDL_rcpt_A.
 DR InterPro: IPR01254; Peptidase_S1_S6.
 DR InterPro: IPR003314; Peptidase_S1A.
 DR InterPro: IPR001190; Srcr_rcpt.
 DR Pfam: PF00089; Trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00192; LDLa_1.
 DR SMART: SM00202; SR_1.
 DR SMART: SM00202; Tryp_spec_1.
 DR PROSITE: PS01209; LDLa_1; 1.
 DR PROSITE: PS50068; LDLa_2; 1.
 DR PROSITE: PS50287; SR_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR HydroLase: Membrane; Protease; Serine protease; Transmembrane.
 DR SEQUENCE 492 AA; 53859 MW; C05B531C8A311C7 CRC64;

Alignment Scores:

| Pred. No.: | Score: | Length: | Matches: |
|------------------------|--------|---------------|----------|
| Percent Similarity: | 768.50 | 492 | 178 |
| Best Local Similarity: | 50.64 | Conservative: | 74 |
| Query Match: | 35.74 | Mismatch: | 195 |
| | 23.14 | Indels: | 51 |
| | | Gaps: | 15 |

US-10-806-370-11 (1-1748) x Q6GTK7_HUMAN (1-492)

QY 288 GGTCTCTATCCGGCAGGTGAT---CATCCGCGAGGTGAGCTCGGTGACAACTCCCA 344
 Db 12 G1YProTyrTyrGluAsnHisGlyTyrGlnProGluAsnPro-----TyrProAlaGln 29
 QY 345 CCAGAGTGT-----ACCTGTGAGGACACACCGAGTGGGGCTGTACCTCC----- 392
 Db 30 ProThrValValProThrValTyrGlnValHisProAlaGlnTyrTyrProSerProVal 49
 QY 393 GATCATCTCTCCGACGTCAGACACACGACGACGAGGCCACGAGGAGAGAGCCAGTACGA 452
 Db 50 ProGlnTyrAlaProArgValLeuThrGlnAlaSerAsnPro-----ValValCys 66
 QY 453 GCCGCGCCAACTTCACTCGTGGGAGGAGGCGGAGGACGACTACCGCTACCGGCGCTGC 512
 Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysAlaLeuCys 86
 QY 513 -----TCTCTCTCATTTGCCCTGG 530
 Db 87 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTyr 106
 QY 531 TGGTTTGGCTCATCTCTTTCAGATTCTGGCAGGCGCACACAGGATCAGTACAGG 590
 Db 107 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125
 QY 591 AGCAGAGGAGAGAGCTGTCCCAACGCGTCTCGTGTGACGCGGTGTGACATGCAAGC 650
 Db 126 -----CysIleAsnProSerAsnTyr--CysAspGlyValSerHisCysProG 141
 QY 651 TGAAGAGTGAAGAGTGGCTGGCTGAGAGTTTGAAGTGGACAAGCTCTGCTTAATCT 710
 Db 141 LysGlyGluAspGluAsnHisGlyValAlaGlyLeuTyrGlyProAsnPheIleLeuGlnValT 161
 QY 711 ACTCTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGACGACGACGATGAGTACTACT 770
 Db 161 YrsrSerGlnArgLysSerThrPheProValCysGlnAspAspTyrPheGlnAsnTyrG 181
 QY 771 CAGAGAAAGCTGCGCAGAGAGTGGGTTTCAGAGTGTCTGACCCGACCAACGAGTGGCC 830
 Db 181 LysArgAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
 QY 831 ACAGGATTTTGCACCAAGCTTTCATCTTGAAGTACGATCACTCCAGC----- 877
 Db 200 LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 220

QY 878 --ATCCAGAAAGCTCCACAGGTCTGA---TCCCTTCCACAGCGGTATATCTCCCTCC 932
 Db 220 spIleTyrLysLeuThrHisSerAspAlaCysSerSerLysAlaValAlaSerLeuA 240
 QY 933 AGTGTCCCATCTGGGACTGAG-----GCCATGACCGCGGATCTGTGGAGGG 983
 Db 240 rGcysIleAlaCysGlyValAlaAsnLeuAsnSerSerArgGlnSerArgIleValGlyG 260
 QY 984 CGCTGCGCTCGATAGCAAGTGGCTTGGCAAGTGTCTGACTTGGCCACCCACA 1043
 Db 260 luserAlaLeuProGlyAlaTyrProTyrGlnValSerLeuHisValGlnAsnValHisV 280
 QY 1044 TCTGTGAGAGCAGCTCATTTGACCCGAGTGGTGTCTCATCTGCGCCACTGCTTCTTCG 1103
 Db 280 alCysGlyLysIleIleThrProGluTyrPheAlaThrAlaAlaHisCysVal---- 298
 QY 1104 TGACCCGAGAGAGTCTCTGAGGCG-----TGAAGGTGTAGCGGGGACCCACA 1154
 Db 299 -----GluLysProLeuAsnAsnProTyrPheIleThrAlaPheAlaGlyIleLeuA 316
 QY 1155 ACTGCACACAGTTCCTGAGGACGCC-----TCCATTGCCAGATCATCTCAACAGA 1208
 Db 316 rGlnSerPheMetPheTyrGlyAlaGlyTyrGlnValGluLysValIleSerHisProA 336
 QY 1209 ATTAACCCGATGAGAGAGCAGTATGACATCGCCCTCATCGGCTGTCCAGCCCTGA 1268
 Db 336 snTyrAspSerLysThrLysAlaAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
 QY 1269 CCTGTCCGCTCATCATCCCTCTGCTTCCCTCCCATGATGACGACACACCTTACCTCA 1328
 Db 356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetLeuGlnProG 376
 QY 1329 ATGACAGCTCTGGATGACAGGCTTTTGGCAAGACGAGGAGACAGATGACACATCCC 1388
 Db 376 lueGlnLeuCysTrpIleSerGlyTyrGlyAlaThrGlnGlu---LysGlyLysThrSerG 395
 QY 1389 CTTCTCTCCGAGAGTCCAGGTCAATTCATTCGACTTCAAGAAATGCAATGACTACTTGG 1448
 Db 395 luvAlaLeuAsnAlaAlaIleValLeuLeuIleGlnThrGlnArgCysAsnSerArgTyrV 415
 QY 1449 TCTATGACAGTACCTTATCCCAAGATGATGTGTGTGGGACCTTGTGGGGGAGAG 1508
 Db 415 alTyrAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnLysAsnValA 435
 QY 1509 ACTCTGCGCAGAGAGAGAGCGGGGGGCTCTGTGTGTGAGACAGAAACGCTGTATCC 1568
 Db 435 epsrCysGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTyrTrpL 455
 QY 1569 TGGCAGGTGTCCACGCTGGGCGACAGGCTGTGGCCAGAGAAACAACTGTGTGTACA 1628
 Db 455 euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTyrArgProGlyValTyrG 475
 QY 1629 CCAAGAGACAGAAAGTCTTCCCTGATTTACAGCAAGATGGAAGGAG 1678
 Db 475 lyaAsnValMetValPheThrAspTrpIleTyrArgGlnMetArgAlaAsp 491

RESULT 13
 Q31UB_HYLSY PRELIMINARY: PRT: 484 AA.
 AC Q31UB_HYLSY
 DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Transmembrane protease serine 2 (Fragment).
 GN Name=TMPRSS2;
 OS Homo sapiens (Homo sapiens) (Homo sapiens).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Hylobatidae; Symphalangus.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Clark N.L., Swanson W.J.;

RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

CC -----

DR EMBL: DQ150504; AA282291.1; -; Genomic DNA.

DR GO: GO:0016021; C: integral to membrane; IEA.

DR GO: GO:0016020; C: membrane; IEA.

DR GO: GO:0008233; F: peptidase activity; IEA.

DR GO: GO:0005044; F: scavenger receptor activity; IEA.

DR GO: GO:0004232; F: serine-type endopeptidase activity; IEA.

DR GO: GO:0006508; P: proteolysis; IEA.

DR InterPro: IPR002172; LDL_rcpt_A.

DR InterPro: IPR001254; Peptidase_S1_S6.

DR InterPro: IPR001314; Peptidase_S1A.

DR InterPro: IPR01190; Src_rcpt.

DR Pfam: PF00089; Ldl_recept_a; 1.

DR Pfam: PF00089; Trypsin; 1.

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR SMART: SM00192; LDLa; 1.

DR SMART: SM00202; SR; 1.

DR SMART: SM00202; TRYP_SPC; 1.

DR PROSITE: PS01209; LDLRA_1; 1.

DR PROSITE: PS00068; LDLRA_2; 1.

DR PROSITE: PS0287; SRCR_2; 1.

DR PROSITE: PS0240; TRYP_SIN_DOM; 1.

DR PROSITE: PS00134; TRYP_SIN_HIS; UNKNOWN_1.

DR PROSITE: PS00135; TRYP_SIN_SER; 1.

KW Protease; Transmembrane.

FT NON_TER 1

FT NON_TER 484

SEQUENCE 484 AA; 53157 MM; 5879BCD0BCE9AE CRC64;

Alignment Scores:

| Pred. No.: | Score: | Length: | Matches: |
|------------------------|--------|---------------|----------|
| 6.41e-35 | 767.50 | 484 | 179 |
| Percent Similarity: | 50.0% | Conservative: | 69 |
| Best Local Similarity: | 36.1% | Mismatches: | 195 |
| Query Match: | 23.0% | Indels: | 53 |
| | | Gaps: | 16 |

US-10-806-370-11 (1-1748) x Q311U8_HYLSY (1-484)

QY 288 GGTCTCATCCGGAGGTCAT---CATCCGCCAGCTGAGCTCGGTGACAACTCCCA 344

DB 7 GlyProTyTYrGluasnHisGlyTYrGlnProGluasnPro-----TyProAlaGln 24

QY 345 CCAGAGTGT-----ACCTGTGAGCAACACCAAGTGGGGCTGTACCCATCC----- 392

DB 25 ProThrValAlaProSerValTYrGluValHisProAlaGlnTYrTYrProSerProVal 44

QY 393 GATCATCTCTCCGAGGTCAACACCAAGCAGGAGGCCAGGAGAGAGCCAGATGACA 452

DB 45 ProGlnTYrAlaProArgValLeuThrGlnAlaSeranPro-----ValValArg 61

QY 453 GCCTGCCCAAGTATTCCTGGCGGAGGCCAGAAAGCAAGCTACCGCTCATCGGGTGCCTG 512

DB 62 ThrGlnProTySerProSerGlyThrValCysThrSerIlyThrIlyValAlaValCys 81

QY 513 -----TCCTCTCATTCGCTGG 530

DB 82 IleThrLeuThrLeuGlyAlaPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp 101

QY 531 TGGTTTGGCTCATCTCTTCGAGTTTGGACAGGCGCACACAGGATCAGGTACAGG 590

DB 102 LysPheMetGlySerIlyCysSerAlaSerGlyIleGluCysAspSerSerGlyIle--- 120

QY 591 AGCAGAGGAGAGCTGTCCCAAGACCGCTGTCCGTGAGGAGGGTGTGATGACGAC 650

DB 121 -----CysIleSerAlaSeranTrp--CysAspGlyValSerHisCysProS 136

QY 651 TGAAGAGTACAGAGCTGGGCTGCGTGAAGTTGACTGGAGCAAGTCTCTGCTTAAATCT 710

DB 136 erGlyGluAspGluAsnGlnCysValAlaArgLeuTYrGlyProAsnPheIleLeuGlnValT 156

QY 711 ACTCTGGCTCTCCCATCATGCTTCCCATCTCTGTACACCAATGGAATGACTCTACT 770

DB 156 yreSerGlnArgIlySerSerTrpHisProValCysGlnAspAspTrpAsnGluasnTYG 176

QY 771 CAGAGAAAGCTTCCGACACACAGCTGGTTTCGAGAGTGTCCACCGACACCGAG---GTTG 827

DB 176 IlyArgAlaAlaCysArgAspMetClyTYrIlyAsnSerPheTYrSerSerGlnIlyLev 196

QY 828 CCCACAGGATTTTCCCAACAGCTTCTCAATCTTGACA-----TACAACTCCACCA 878

DB 196 AlaAspSerGlyAlaThrSerPheMetIlyLeuAnThrSerAlaArgAsnValAspI 216

QY 879 TCCAGAAAGCTTCCACAGCTTGAA---TGCCCTTCCACAGCGTATATCTCTCCAGT 935

DB 216 IeTYrIlyValSerIlyThrIlySerAspAlaCysSerSerIlyValAlaValSerLeuArgC 236

QY 936 GTTCCCACTGGGAGGAGG-----GCCATGACCGGCGGAGTGTGGAGGAGGCGC 986

DB 236 yslIleAlaCysGlyValAsnLeuAsnSerSerArgIlySerArgIlyValGlyIlyGly 256

QY 987 TGGCTTGGATGACAGTGGCTTGGCAAGTGAAGTGTGCACTTGGACACCACTT 1046

DB 256 erAlaArgLeuGlyAlaTrpProTrpGlnValSerLeuHisValGlnAsnValIleValC 276

QY 1047 GTGAGGACAGCTCATTTGACGCCCGAGTGGTCTCATCGCCGACCTGCTTCTTCGTA 1106

DB 276 yslGlySerIleIleThrProGlnTrpIleValIleAlaAlaHisCysVal----- 293

QY 1107 CCGGAGAGAGTCTGTGAGGAGC-----TGAAGTGTATCGGCGGACACGAAAC 1157

DB 294 -----GluTyProLeuAsnAsnProTrpHisIleTrpHisAlaPheAlaGlyIle-----L 310

QY 1158 TGACACAG-----TTGCTTGGAGCAGCTCATTTGGCAGATCATCATCA 1205

DB 310 euArgIlySerPheMetPheTYrGlnSerAlaHisGlnValGlyValIleSerHisP 330

QY 1206 GCATTTACCCGATGAGAGAGACATATGACATCGCCCTCATTCGCGCTGTCCAGCCCC 1265

DB 330 roAsnTYrAspSerIlyThrIlyAsnAsnAspIleAlaLeuMetIlyLeuGlnThrProL 350

QY 1266 TGACCTGTCCGCTCATCATCAACCTGTGCTTCCCTCCCATGATGACAGACACTTTAGCC 1325

DB 350 eutrPheAsnAspPheValIlyProValCysLeuProAsnProGlyLeuMetLeuGlnP 370

QY 1326 TCAATGAGACTGTGTGATCAGAGCTTTGGCAAGACAGGAGACAGATGACAAAGAT 1385

DB 370 roGlnGlnHisCysTrpIleSerGlyTrpGlyAlaThrGlnGln---LysGlyIlyThrS 389

QY 1386 CCCCCTTCTCCGAGAGTGCATATTCATGCACTTCACAGAAATGCATATGACTACT 1445

DB 389 erAspValIleAsnAlaAlaMetValArgLeuIleGlnThrGlnArgCysAsnSerArgT 409

QY 1446 TGGCTATGACATGACTTATCCATCCACAGATGATGTGTGCGGGAGCTTGTGGGGGCA 1505

DB 409 yValIlyAspAsnLeuIleThrProAlaMetIlyCysAlaGlyPheLeuGlnIlyThrV 429

QY 1506 GAGACTCTCCGACAGGAGACAGCGGGGAGCTTGTGTGTGACAGACAAACCGCTGT 1565

DB 429 alaAspSerCysGlnGlyAspSerGlyIlyProLeuValIleThrSerIlyAspAsnValTrp 449

QY 1566 ACCTGAGAGTGTACACAGCTGGGGGACAGGCTGTGGCCACAGAAACAACTGTGTGT 1625

DB 449 rPheIleGlyAspTrpHisSerTrpGlySerGlyCysAlaIlyValIlyArgProGlyValT 469

QY 1626 ACACCAAGTACAGAGATCTTCTCCGATTTACACAGATG 1669

DB 469 yrgIlyAsnValIleValPheThrAspTrpIleTYrArgGlnMet 483

RESULT 14

TPMS2_HUMAN STANDARD; PRT; 492 AA.

ID TPMS2_HUMAN

AC 015393; Q9BXN1;
 DT 15-JUN-1998, integrated into UniProtKB/Swiss-Prot.
 DT 18-OCT-2001, sequence version 2.
 DT 07-MAR-2006, entry version 54.
 DE Transmembrane protease, serine 2 precursor (EC 3.4.21.-) [Contains:
 DE Transmembrane protease, serine 2 non-catalytic chain; Transmembrane
 DE protease, serine 2 catalytic chain].
 GN Name=TMPRSS2; Synonyms=PRSS10;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 NCBI_TaxID=9606;
 RN NUCLEOTIDE SEQUENCE [MRNA].
 RX MEDLINE=97468144; PubMed=9325052; DOI=10.1006/geno.1997.4845;
 RA Polonri-Giacchino A., Chen H., Peltech M.C., Rosseter C.,
 RA Antonarakis S.E.;
 RT "Cloning of the TMPRSS2 gene, which encodes a novel serine protease
 RT with transmembrane, IDRA, and SRCR domains and maps to 21q22.3.";
 RL Genomics 44:309-320(1997).
 [2]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RX MEDLINE=21305069; PubMed=11414763; DOI=10.1006/geno.2001.6551;
 RA Teng D.-H., Chen Y., Lian L., Ha P.C., Tavtigian S.V., Wong A.K.C.;
 RT "Mutation analyses of 268 candidate genes in human tumor cell lines.";
 RL Genomics 74:352-364(2001).
 [3]
 RN NUCLEOTIDE SEQUENCE [MRNA] AND MUTAGENESIS.
 RX MEDLINE=21139112; PubMed=11245484;
 RA Alter D.E.H., Vivanco I., Hubert R.S., Kuo J., Chen E., Saffran D.C.,
 RA Raitano A.B., Jakobovits A.;
 RT "Catalytic cleavage of the androgen-regulated TMPRSS2 protease results
 RT in its secretion by prostate and prostate cancer epithelia.";
 RL Cancer Res. 61:1686-1692(2001).
 [4]
 RN TISSUE SPECIFICITY.
 RX MEDLINE=21104370; PubMed=11169526;
 RA DOI=10.1002/1096-9896(2000)9999:9999::AID-PATH743>3.0.CO;2-T;
 RA Vazirani M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Vahko P.T.;
 RT "Expression of transmembrane serine protease TMPRSS2 in mouse and
 RT human tissues.";
 RL J. Pathol. 193:134-140(2001).
 [5]
 RL SUBCELLULAR LOCATION: Type II membrane protein. Activated by
 [6]
 CC cleavage and secreted.
 CC TISSUE SPECIFICITY: Expressed strongly in small intestine. Also
 CC expressed in prostate, colon, stomach and salivary gland.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family.
 CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -1- SIMILARITY: Contains 1 SRCR domain.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: U75329; AAC51784.1; -; mRNA.
 CC EMBL: AF123453; AAD3117.1; -; mRNA.
 CC EMBL: AF270487; AAK29280.1; -; mRNA.
 CC HSP: P00760; IEZX.
 CC MEROPS: S01.247; -.
 CC Ensembl: ENSG00000184012; Homo sapiens.
 CC HGNC: HGNC:11876; TMPRSS2.
 CC MIM: 602060; gene.
 CC GO: GO:0005887; C:integral to plasma membrane; TAS.
 CC GO: GO:0008236; F:serine-type peptidase activity; TAS.
 CC InterPro: IPR002172; LDL_rcpt_A.
 CC InterPro: IPR001254; Peptidase_S1_S6.
 CC InterPro: IPR001314; Peptidase_S1A.
 CC InterPro: IPR001190; Srcr_rcpt.
 CC Pfam: PF00089; Trypsin_1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00192; LDLA; 1.
 CC SMART: SM00202; SR; 1.

DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS01209; LDLRA_1; 1.
 DR PROSITE; PS00068; LDLRA_2; 1.
 DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
 DR PROSITE; PS00287; SRCR_2; 1.
 DR PROSITE; PS02440; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Hydrolase; Membrane; Polymorphism; Protease;
 KW Serine protease; Signal-anchor; Transmembrane; Zymogen.
 FT CHAIN 1 255
 FT Transmembrane protease, serine 2 non-
 FT catalytic chain.
 FT /Frid=PRO_0000027855.
 FT Transmembrane protease, serine 2
 FT catalytic chain.
 FT /Frid=PRO_0000027856.
 FT Cytoplasmic (potential).
 FT Signal-anchor for type II membrane
 FT protein (potential).
 FT Extracellular (potential).
 FT LDL-receptor class A.
 FT SRCR.
 FT Peptidase S1.
 FT Charge relay system (By similarity).
 FT Charge relay system (By similarity).
 FT Charge relay system (By similarity).
 FT Cleavage (potential).
 FT Cleavage (potential).
 FT N-linked (GlcNAc...) (potential).
 FT N-linked (GlcNAc...) (potential).
 FT By similarity.
 FT By similarity.
 FT By similarity.
 FT By similarity.
 FT By similarity.
 FT By similarity.
 FT Interspersed (between non-catalytic and
 FT catalytic chains) (By similarity).
 FT By similarity.
 FT By similarity.
 FT By similarity.
 FT By similarity.
 FT VARIANT 449 449
 FT MUTAGEN 255 255
 FT MUTAGEN 441 441
 FT CONFLICT 160 160
 FT CONFLICT 242 242
 FT CONFLICT 329 329
 FT CONFLICT 489 491
 FT SEQUENCE 492 AA; 53891 MW; CAB44FD174A8076B CRC64;
 SO
 Alignment Scores:
 Pred. No.: 7,31e-35 Length: 492
 Score: 766.50 Matches: 178
 Percent Similarity: 50.6% Conservative: 74
 Best Local Similarity: 35.7% Mismatches: 195
 Query Match: 23.0% Indels: 51
 Gaps: 15
 US-10-806-370-11 (1-1748) x TMPRSS2_HUMAN (1-492)
 QY 288 GGTCTCATCGGAGAGTCA---CATCCGCGAGCTGAGCTGATCAACATCTCCCA 344
 DB |||||
 Db 12 GlyProTyrTyrGluAsnHsGlyTyrGlnProTyrLysPro-----TyrProAlaGln 29
 QY 345 CCAGAGTGT-----ACCTGTGAGACACACAGGAGGAGGCTGATCCATCC----- 392
 DB |||||
 Db 30 ProThrValValProThrValTyrGluValHisProAlaGlnTyrTyrProSerProVal 49
 QY 393 GATCATCTCTGCGAGCTGACGACACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 452
 DB |||||
 Db 50 ProGlnTyrAlaProAlaGValLeuThrGlnAlaSerAsnPro-----ValValCys 66
 QY 453 GCTGCCCAAGTAACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512
 DB |||||
 Db 453 GCTGCCCAAGTAACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512

```

Db      67  ThrGlnProlyserProserGlyThrValCysThrSerlyThrLysAlaLeuGlyCys  86
QY      513  -----TCTCTCTCATTCGCTCG  530
Db      87  IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp  106
QY      531  TGGTGGTCATCATCTCTTCAGATCTTGACAGGGGACACACAGGGATGATGATACAG  590
Db      107  LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr---  125
QY      591  AGCAGAGGAGAGGCTGTCCCAAGCAGCGTGTCTGCTGACGGGGGTGTGATGACACG  650
Db      126  -----CysIleAsnProserAsnTrp--CysAspGlyValSerHisCysProG  141
QY      651  TGAAGATGACAGAGCTGGGCTGCTGCTGCTTGAAGCTGGGACAAAGTCTCTGTTAAATCT  710
Db      141  LysGlyLysAspGlyAsnAsnArgCysValArgLeuTrpGlyProAsnPheIleLeuGlnMet  161
QY      711  ACTCTGGGCTCTCCATCATGAGCTGCTCCATCTGTAGACAGACATGATGATCTTACT  770
Db      161  YrSerSerGlnArgLysSerTrpHisProValCysGlnAspAspTrpAsnGluAsnTrpG  181
QY      771  CAGAGAAAGCTGACAGAGCTGGGTTTCAGAGTGTCTACCGGACCAACGAGGTTGCC  830
Db      181  LysArgAlaAlaCysArgAspMetGlyTrpLysAsnAsnPheTyrSerSerGln---GlyI  200
QY      831  ACAGGATTTTCCACAGCTTCTCATCTTGAATACATCCAC-----  877
Db      200  LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA  220
QY      878  --ATCCAGAAAGGCTCCACAGCTGAA--TGCCTTCCACGGGATATCTCCCTCC  932
Db      220  spIleTrpLysLeuLeuTrpHisSerThrAlaCysSerSerLysAlaValAlaSerLeu  240
QY      933  AGTGTTCACATGCGGACTGAG-----GCCATGACCGGCGGATCTGGGAGGAG  983
Db      240  rGysIleAlaCysGlyValAlaAsnLeuAsnSerSerArgGlnSerArgIleValGlyG  260
QY      984  CGTGGCTCGGATGACAGAGTGGCTTGGAAGTGTGACCTTCCGACCCACCA  1043
Db      260  LysSerAlaLeuProGlyAlaTrpProTrpGlnValSerLeuHisValGlnAsnValHis  280
QY      1044  TCTGTGAGGAGACCTTATTAACCCAGTGGGTGTCTACCTGCGCCACAGCTTCTTG  1103
Db      280  alCysGlyGlySerIleThrProGluTrpIleValThrAlaAlaHisCysVal----  298
QY      1104  TGACCCGGAGAGAGTCTTGAGGAGC-----TGGAAGGTGTACCGCGGACCA  1154
Db      299  -----GluLysProLeuAsnAsnProTrpHisTrpThrAlaPheAlaGlyIleLeu  316
QY      1155  ACCTGACACAGTTCCTGAGGAGC-----TCCATTGCCAGATCATCAACAGCA  1208
Db      316  rGlnSerPheMetPheTrpGlyAlaGlyTrpGlnValGlnValIleSerHisProA  336
QY      1209  ATTACACCGATGAGAGAGACATATGACATGCGCTTCATGCGCTGTCACAGCCCTGA  1268
Db      336  enTrpAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeu  356
QY      1269  CCCTGCGCTCATCATCCACCTGCTGCTCCCATCATGACAGACAGACCTTTAGCTCA  1328
Db      356  hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetCysLeuGlnProG  376
QY      1329  ATGAGACCTGCTGATTCACAGGCTTTGGCAAGACAGGAGAGACAGATGACAGACATCC  1388
Db      376  LysGlnLeuCysTrpIleSerGlyTrpGlyAlaTrpGlnGlnG---LysGlyLysThrSerG  395
QY      1389  CCTTCTCCGAGAGGTGACATGATCATGATGATGATGATGATGATGATGATGATGATG  1448
Db      395  LysValLeuAsnAlaAlaLysValLeuLeuIleGluThrGlnArgGlySerHisArgTrp  415
QY      1449  TCTATGACAGTTTACCTTACCCCAAGATGATGTGTGCTGGGAGCTTGTGTGGGGCAGAG  1508
Db      415  alTrpAspAsnLeuLeuThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA  435

```

```

QY      1509  ACTCTGCGCAGGGAGAGACAGCGGGGCTCTGTGTGAGACAGAACACCGCTGTACC  1568
Db      435  spSerCysGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTrpTrp  455
QY      1569  TGGCAGGTGTACACAGCTGGGGGACAGCTGTGGCCAGAGAAACAACTGTGTGTACA  1628
Db      455  euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTrpArgProGlyValTrpG  475
QY      1629  CCAAAGTACAGAAAGTTCTTCCCTGATTTACAGCAAGATGAGAGCGCAG  1678
Db      475  LysAsnValMetValPheThrAspTrpIleTrpArgGlnMetArgAlaAsp  491

RESULT 15
ID      096T73_HUMAN      PRELIMINARY;      PRT;      492 AA.
AC      096T73;
DT      01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT      01-DEC-2001, sequence version 1.
DT      21-FEB-2006, entry version 26.
DE      Epithelialin.
GN      Name=TPRSS2;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC      Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21223025; PubMed=11322890;
RA      Jacquinet E., Rao N.V., Rao G.V., Wang Z., Albertine K.H.,
RA      Hoidal J.R.;
RT      "Cloning and characterization of the cDNA and gene for human
RT      epithelialin.";
RL      Eur. J. Biochem. 268:2687-2693(2001).
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC      EMBL; AF329454; AAK53559.1; -; mRNA.
DR      HSSP; P00766; 1CHG.
DR      Ensembl; ENSG00000184012; Homo sapiens.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005044; F:scavenger receptor activity; IEA.
DR      GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR      GO; GO:0006508; P:proteolysis; IEA.
DR      InterPro; IPR002172; IDL_rcpt_A.
DR      InterPro; IPR001254; Peptidase_S1_S6.
DR      InterPro; IPR001314; Peptidase_S1A.
DR      InterPro; IPR001190; Src_rcpt.
DR      Pfam; PF00089; Tyrosin_1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      SMART; SM00192; LDla; 1.
DR      SMART; SM00202; SR; 1.
DR      SMART; SM00020; TYP_SPC; 1.
DR      PROSITE; PS01209; LDla_1; 1.
DR      PROSITE; PS50068; LDla_2; 1.
DR      PROSITE; PS50287; SRC_2; 1.
DR      PROSITE; PS50240; TRYPsin_DOM; 1.
DR      PROSITE; PS00134; TRYPsin_His; UNKNOWN_1.
DR      PROSITE; PS00135; TRYPsin_SER; 1.
KW      Hydrolyase; Membrane; Protease; Serine protease; Transmembrane.
SQ      SEQUENCE 492 AA; 5363 MW; 3ABA755B276DADF CRC64;

Alignment Scores:
Pred. No.:      7,31e-35      Length:      492
Score:          766.50      Matches:      178
Percent Similarity: 50.6%      Conservative: 74
Best Local Similarity: 35.7%      Mismatches:  195
Query Match:      23.0%      Indels:      51
DB:              2          Gaps:      15

```

US-10-806-370-11 (1-1748) x 096T73_HUMAN (1-492)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 01:52:32 ; Search time 12.6 Seconds
(without alignments)
3642.941 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 3333
Sequence: 1 cccagagaccatgagagaggg.....ggtctgtgtgactcagagaa 1748

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 650591 seqs, 87530628 residues
Total number of hits satisfying chosen parameters: 1301182

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p.model -DEV=xip
-Q=/abs/ABSSWEB.spool/US10806370/runat.15092006.105634.10230/app.query.fasta_1
-DB=Issued Patents_AA -OPMT=fastaan -SUFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02p -USER=US10806370 -CGEN 1.1.84 @runat.15092006.105634.10230
-NCPU=6 -ICPU=3 -NO_MMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGIOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_AA:
1: /EMC_Celerra_SIDS3/prodata/2/1aa/5 COMB.pep.*
2: /EMC_Celerra_SIDS3/prodata/2/1aa/6 COMB.pep.*
3: /EMC_Celerra_SIDS3/prodata/2/1aa/7 COMB.pep.*
4: /EMC_Celerra_SIDS3/prodata/2/1aa/H COMB.pep.*
5: /EMC_Celerra_SIDS3/prodata/2/1aa/ECTUS COMB.pep.*
6: /EMC_Celerra_SIDS3/prodata/2/1aa/RE COMB.pep.*
7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfill1est.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------------------------|
| 1 | 2999 | 90.0 | 562 | 2 | US-09-879-792-12 Sequence 1, Appl |
| 2 | 2291.5 | 68.8 | 477 | 2 | US-10-177-661-2 Sequence 2, Appl |
| 3 | 2108 | 63.2 | 446 | 2 | US-10-177-661-4 Sequence 4, Appl |
| 4 | 969.5 | 29.1 | 418 | 2 | US-10-177-661-6 Sequence 6, Appl |
| 5 | 769.5 | 23.1 | 492 | 2 | US-09-759-143-932 Sequence 932, App |
| 6 | 769.5 | 23.1 | 492 | 2 | US-10-012-996-932 Sequence 932, App |
| 7 | 769.5 | 23.1 | 492 | 3 | US-10-144-678A-932 Sequence 932, App |
| 8 | 768.5 | 23.1 | 492 | 2 | US-09-342-749-2 Sequence 2, Appl |
| 9 | 768.5 | 23.1 | 492 | 2 | US-09-691-840-2 Sequence 2, Appl |
| 10 | 768.5 | 23.1 | 510 | 2 | US-09-949-016-11074 Sequence 11074, A |
| 11 | 764.5 | 22.9 | 492 | 2 | US-09-685-166A-895 Sequence 895, App |

| | | | | | | |
|----|-------|------|-----|---|---------------------|-------------------|
| 12 | 764.5 | 22.9 | 492 | 2 | US-09-879-792-14 | Sequence 14, Appl |
| 13 | 764.5 | 22.9 | 492 | 2 | US-09-679-442-895 | Sequence 895, App |
| 14 | 764.5 | 22.9 | 492 | 2 | US-09-759-143-895 | Sequence 895, App |
| 15 | 764.5 | 22.9 | 492 | 2 | US-10-012-896-895 | Sequence 895, App |
| 16 | 764.5 | 22.9 | 492 | 3 | US-10-144-678A-895 | Sequence 934, App |
| 17 | 766.5 | 22.1 | 393 | 3 | US-09-759-143-934 | Sequence 934, App |
| 18 | 736.5 | 22.1 | 393 | 3 | US-10-012-896-934 | Sequence 934, App |
| 19 | 736.5 | 22.1 | 393 | 3 | US-10-144-678A-934 | Sequence 934, App |
| 20 | 736.5 | 22.1 | 521 | 2 | US-09-949-016-11081 | Sequence 11081, A |
| 21 | 736.5 | 22.1 | 521 | 2 | US-09-949-016-11082 | Sequence 11082, A |
| 22 | 736.5 | 22.1 | 521 | 2 | US-09-949-016-11083 | Sequence 11083, A |
| 23 | 729.5 | 21.9 | 453 | 2 | US-09-999-833A-69 | Sequence 69, Appl |
| 24 | 729.5 | 21.9 | 453 | 2 | US-10-020-445A-69 | Sequence 69, Appl |
| 25 | 729.5 | 21.9 | 453 | 2 | US-09-978-189-69 | Sequence 69, Appl |
| 26 | 729.5 | 21.9 | 453 | 2 | US-10-017-085A-69 | Sequence 69, Appl |
| 27 | 729.5 | 21.9 | 453 | 3 | US-10-145-129A-69 | Sequence 69, Appl |
| 28 | 729.5 | 21.9 | 453 | 3 | US-10-013-929A-69 | Sequence 69, Appl |
| 29 | 729.5 | 21.9 | 453 | 3 | US-10-013-917A-69 | Sequence 69, Appl |
| 30 | 726 | 21.8 | 454 | 2 | US-09-518-046-2 | Sequence 2, Appl |
| 31 | 726 | 21.8 | 454 | 2 | US-09-550-371-2 | Sequence 2, Appl |
| 32 | 704 | 21.1 | 423 | 2 | US-09-656-002-2 | Sequence 2, Appl |
| 33 | 697.5 | 20.9 | 435 | 2 | US-09-607-745-2 | Sequence 2, Appl |
| 34 | 697.5 | 20.9 | 435 | 3 | US-10-030-688-2 | Sequence 2, Appl |
| 35 | 697.5 | 20.9 | 437 | 2 | US-09-851-588-8 | Sequence 8, Appl |
| 36 | 696.5 | 20.9 | 435 | 2 | US-09-008-271A-6 | Sequence 6, Appl |
| 37 | 696.5 | 20.9 | 435 | 2 | US-09-968-415-6 | Sequence 6, Appl |
| 38 | 695.5 | 20.9 | 406 | 2 | US-09-851-588-6 | Sequence 6, Appl |
| 39 | 695 | 20.9 | 492 | 3 | US-10-030-688-4 | Sequence 4, Appl |
| 40 | 694 | 20.8 | 432 | 2 | US-10-012-231A-275 | Sequence 275, App |
| 41 | 694 | 20.8 | 432 | 2 | US-10-015-389A-275 | Sequence 275, App |
| 42 | 694 | 20.8 | 432 | 2 | US-10-006-768A-275 | Sequence 275, App |
| 43 | 694 | 20.8 | 432 | 2 | US-10-015-671A-275 | Sequence 275, App |
| 44 | 694 | 20.8 | 432 | 2 | US-10-015-393A-275 | Sequence 275, App |
| 45 | 694 | 20.8 | 432 | 2 | US-10-011-833A-275 | Sequence 275, App |

ALIGNMENTS

RESULT 1
US-09-879-792-12
; Sequence 12, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-MO)
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-879-792-12
Alignment Scores:
Pred. No.: 2,11e-218
Score: 2999.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 90.0%
Length: 562
Matches: 562
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

US-10-806-370-11 (1-1748) x US-09-879-792-12 (1-562)
QY 11 ATGAGAGGAGCAGACACGGGATGATCTTCACAGAAAGAACCTTCAGTGGACATCT 70
Db 1 MetGluArgAspSerHisGlyAsnHisLaserProAlaHisGlnProSerAlaGlyAlaSer 20
QY 71 CCAGCCCAAGGATCTCCAGCTGGAGACATTCACAGCCCGGAGCATCTCCAGCCAGGACATCT 130
Db 21 ProAlaGlnHisLaserProAlaGlyThrProProGlyArgAlaSerProAlaGlnHisLaser 40
QY 131 CCAGCCCAAGGATCTCCAGCTGGAGACATTCAGCCCGGAGCATCTCCAGCCAGGACATCT 190
Db 41 ProAlaGlnHisLaserProAlaGlyThrProProGlyArgAlaSerProAlaGlnHisLaser 60
QY 191 CCAGCTGGATACATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCAGGACATCT 250
Db 61 ProAlaGlyThrProProGlyArgAlaSerProAlaGlyThrProAlaGlnHisLaser 80
QY 251 CCAGCCCGGAGCATCTCCAGCTGGAGACATTCAGCTTCAGCTTCATCCGAGGATCACTCA 310
Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerHisSerSerSerGlyArgSerSer 100
QY 311 TCAGCCAGGATGAGCTGGGTACAACTCCCAACAGAGTGTACTTGTAGACAA 370
Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCAGTGGGGGCTGACCATCCAGTATCTCTCCAGCTGGAGACAGCAAGAGGAGC 430
Db 121 ProValGlyAlaValProIleArgSerSerProAlaHisSerAlaProAlaThrArgAla 140
QY 431 ACCAGGAGAGCCAGGATGACAGCTCCGACCAAGTTCACCTGGCGGAGAGCCAGAGCAG 490
Db 141 ThrArgLysSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlnGln 160
QY 491 CTACCGCTCATTCGGGTGGGTCTCTCTCTCATTCGCTGGGTGGGTCTCTCTCTCTCTCTCT 550
Db 161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValAlaSerLeuIleIleLeu 180
QY 551 TTCAGTCTTGGAGGGGACACAGAGGATGACAGGAGAGAGAGAGAGAGTGTCCC 610
Db 181 PheGlnPheTrpGlnGlnGlnHisThrGlyIleArgGlyTyrGlnGlnArgGlnSerCysPro 200
QY 611 AAGCAGCTGTCTTCGCTGACAGGAGTGGTGGTACCTGCAAGTGAAGTGAAGTGAAGTGGC 670
Db 201 LysHisAlaValAlaGlyCysArgGlyValValAlaSerCysLysLeuLysSerArgGlnLeuGly 220
QY 671 TGCCTGAGGTTTGACTGGAGACAGTCTCTGTTAAATCTACTGGGTCTCCCATACAG 730
Db 221 CysValArgPheArgTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHisGln 240
QY 731 TGGCTTCCCATCTGTAGACAGCACTGGAATGACTCTACTAGAGAAAGAACTGACAGAG 790
Db 241 TrpLeuProIleCysSerSerHisTrpAsnLysSerTyrSerGlnLysThrCysGlnGln 260
QY 791 CTGGGTTTCGAGATGCTCACCGGACAAACCGAGTTGGCCACAGGATTTTCCCAAGAGC 850
Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGlnValAlaHisArgAspPheAlaHisSer 280
QY 851 TTCTCAATCTTGAATACAACTCCACCACTCAGAAAGCTCCCAAGTCTGAATGCCCT 910
Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnLysSerLeuHisArgSerGlnCysArgPro 300
QY 911 TCCAGAGCGATATCTCCCTCCAGGTTCACCACTGCGGAGTGAAGGAGCATGACCGGAGCG 970
Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuAlaGlyAlaMetThrGlyArg 320
QY 971 ATCGTGGAGAGGAGCTGAGCTCGAATAGCAAGTGGCTTGGCAAGTGAAGTCTGCATTC 1030
Db 321 IleValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHisPhe 340
QY 1031 GGACCAACCCCATCTGTGAGGAGCAGCTCATTTAGCCCGAGTGGGTCTACTGGCCGCC 1090

```

```

Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
QY 1091 CACTGCTTCTTCGACCCCGGAGAGAGTCTCGAGGGCTGAAAGTGTACGCGGAGCC 1150
Db 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
QY 1151 AGCAACCTGCACCAATTGCTTGAAGCAGCTCCATTCGCCAGATCATCATCAAGCAAT 1210
Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleAsnSerAsn 400
QY 1211 TACACGATAGAGAGAGCACTATGACATGACCTTCATGGGGCTGTCAGCCCTGACC 1270
Db 401 TyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY 1271 CTGCTCCGCTACATCCACCTGCTTGGCTCCCATGATGAGACAGACTTTAGCTCAAT 1330
Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY 1331 GAGACTGTGATACAGGCTTTGGCAGAGCAAGGAGACAGATGACAAATGACATCCCC 1390
Db 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
QY 1391 TTCTCCGGAGAGTGCAGGTCATTCATGACTTCAGAAATGCCATGACTTGTC 1450
Db 461 PheLeuArgGlnValGlnValAlaMetLeuIleAspPheLysCysAsnAspTyrLeuVal 480
QY 1451 TATGACAGTACCTTACCCCAAGATGATGTGCTGGGAGCTTCGTGGGGGAGAGAC 1510
Db 481 TyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeuAlaGlyArgAsp 500
QY 1511 TCCTGCCAGGAGACAGCGGGGAGCTCTTGTCTGTGAGCAGAAACACCTGTGTACAC 1570
Db 501 SerCysGlnGlnLysAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY 1571 GCAAGTGTACACAGCTGGGAGCAGAGCTGCGCAGAGAAACAACTGTGTGTACAC 1630
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnHisGlnLysProGlyValTyrThr 540
QY 1631 AAAGTGACAGAGTCTTCCCTGATTTACAGCAAGTGAAGAGCAGTGCATTCAGA 1690
Db 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetGlnSerGlnValArgPheArg 560
QY 1691 AATTC 1696
Db 561 LysSer 562

RESULT 2
US-10-177-661-2
/ Sequence 2, Application US/10177661
/ Patent No. 6794173
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, Dirk M.
/ TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
/ FILE REFERENCE: 3256-A
/ CURRENT APPLICATION NUMBER: US/10/177,661
/ PRIOR APPLICATION NUMBER: 2002-06-20
/ PRIOR FILING DATE: 2001-06-20
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 477
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-177-661-2

Alignment Scores:
Pred. No.: 5,93e-165 Length: 477
Score: 2291.50 Matches: 424
Percent Similarity: 96.6% Conservative: 4
Best Local Similarity: 95.7% Mismatches: 12
Query Match: 68.8% Indels: 3

```

DB: 2 Gaps: 1
US-10-806-370-11 (1-1748) x US-10-177-661-2 (1-477)
QY 368 ACAACAGTGGGGCTGTACCCATCCATCTCTCCAGGTCAGACCAAGCAACCAGG 427
DB 38 SerProthrinleuSerIleSerSerAspSerSerProIaIaIaIa 57
QY 428 GCCACACAGGAGAGCCAGGTACGAGCTCCCAAGTTCACCTGGCGGAGGGCCAGAG 487
DB 58 LeuLeuLeu-----GlyThrSerLeuProIySerPheThrIrpAsgIuGlnIly 74
QY 488 CAGTACCGGTCTCATCGGGGTGCTGCTCTCTCATTCCTCGGAGGTTTGGCTCATCTC 547
DB 75 GlnLeuProIeuIleGlyCysValLeuLeuLeuIleAlaLeuValIaSerIleIle 94
QY 548 CTCTTCAGTTCGTGACAGGAGCCACAGAGGATCAGGTACAGAGCAGAGAGAGCTGT 607
DB 95 LeuPheGlnPheIrpGlnIlyIleIsthrGlyIleArgIlyIleGlnIlyIleGlnIly 114
QY 608 CCCAAGCAGCCTGTCTGCTGTGACGGGGGTGTGAGTCTGCAAGTGAAGTGAAGCTG 667
DB 115 ProIyHisIleAlaValArgCysAspGlyValIaAspCysIlySerIySerAspGlnLeu 134
QY 668 GGCTGGGTGAGTGTGACGAGGCAAGCTCTGCTTAAATCTACTGGGCTCTCCCAT 727
DB 135 GlyCysValArgPheAspIrpAspIlySerLeuLeuIlyIleIySerGlySerHis 154
QY 728 CAGTGGCTTCCTCATCTGTAGCAGCAACTGGAATGACTCTTACTCAGAGAACTGCGCAG 787
DB 155 GlnIrpLeuProIleCysSerSerAspIrpAspAspSerIySerGlnIlyIsthrCysGln 174
QY 788 CAGCTGGGTTTGCAGAGTGTCTACCCGAGCAACGAGGTTTCCACAGGATTTTGCAC 847
DB 175 GlnLeuGlyPheGlnSerAlaHisArgIsthrThrGlnValAlaHisArgAspPheAlaAsn 194
QY 848 AGCTTCTCATCTTGATGATCAACTCCACCATCCAGAAAGCCCTCCAGAGTGAATGC 907
DB 195 SerPheSerIleLeuArgIlyAspSerThrIleGlnIlySerLeuHisArgSerGlnCys 214
QY 908 CCTTCCAGGAGTATATCTCCCTCCAGTGTTCCTCAGTCCAGTCCAGTCCAGG 967
DB 215 ProSerGlnArgIlyIleSerIleGlnIlyCysSerHisGlyLeuArgAlaMetThrGly 234
QY 968 CGGATCGTGGAGGAGGCGCTGCGCTCGGATGACAGTGGCTTGGCAAGTGTGAC 1027
DB 235 ArgIleValGlyIleAlaLeuAlaSerAspSerIyIrpProIrpGlnIlyIleSerHis 254
QY 1028 TTGGGACCAACCACTGTGTGAGGAGCGCTCATTTAGCGCCAGTGGGCTCAGTGGC 1087
DB 255 PheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnIrpValLeuThrAla 274
QY 1088 GCCCAGTCTCTTCTGTGACCCGAGAGAAAGTCTGAGGGCTGGAAGTGTAGCGCGG 1147
DB 275 AlaHisCysPheHehValThrArgIlyIlyValLeuGlnIlyIrpIyValIyAlaGly 294
QY 1148 ACCAGCAACTGCAACCACTTGGCTGAGCAGCTTCCATTCGCCAGATCATCATCAAGC 1207
DB 295 ThrSerAsnLeuHisGlnLeuProGlnAlaIaSerIleAlaGlnIleIleIleAsnSer 314
QY 1208 AATTACACGATGAGAGGAGCACTATGACATGCGCTCATGGGCTGTCAGAGCCCTG 1267
DB 315 AsnIyThrAspGlnIyAspAspIyAspIleAlaLeuMetArgLeuSerIySproLeu 334
QY 1268 ACCGTGCTCGCTCATCATCAACCTGCTTGGCTCCCATGATGAGCAAGCTTTAGCTTC 1327
DB 335 ThrLeuSerAlaHisIleHisIleSproIaCysLeuProMetHisGlyIsthrPheSerLeu 354
QY 1328 AATGAGACCTGCTGATCAGAGGTTTGGCAAGACCGAGGAGACATGACAAAGATCC 1387
DB 355 AsnGlnIsthrIyIsthrIyIsthrIyPheGlyIlyIsthrArgIsthrAspIyIsthrSer 374
QY 1388 CCTTCTCTCGGAGGAGTGAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTCTG 1447

DB 375 ProPheLeuArgGlyValGlnValaLeuLeuIleAspPheIySerCysAsnAspIyIrpLeu 394
QY 1448 GTCTATGACAGTTACTTACCCCAAGAGATGATGTGTGCTGGGACCTTGTGGGGCAGA 1507
DB 395 ValIyAspSerIyIrpLeuThrProArgMetCysAlaGlyAspLeuIyGlyIyArg 414
QY 1508 GACTCTGCGAGGAGACAGCGGGGGCTCTTGTCTGAGAGAGAAACCGCTGGTAC 1567
DB 415 AspSerCysGlnIlyAspSerGlyIyProLeuValCysGlnIlyAsnAsnArgIrpIy 434
QY 1568 CTGGCAGTGTCTACAGCTGGGAGCAGAGCTGTGGCCAGAGAAACCTGTGTGTAC 1627
DB 435 LeuAlaGlyValIsthrSerIrpGlyIsthrGlyCysGlyGlnArgAsnIySproGlyValIy 454
QY 1628 ACCAAGTGAAGAGTTCCTCTGATTTACAGCAAGTGAAGAGCAGGTCGATTC 1687
DB 455 ThrIyValIsthrGlnValLeuProIrpIleIySerIySerMetGlnSerGlnIyArgPhe 474
QY 1688 AGAAATCC 1696
DB 475 ArgIySer 477

RESULT 3
US-10-177-661-4
; Sequence 4, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Varca, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177, 661
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/299, 606
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-661-4

Alignment Scores:
Pred. No.: 4,29e-151 Length: 446
Score: 2108.00 Matches: 393
Percent Similarity: 89.6% Conservative: 4
Best Local Similarity: 88.7% Mismatches: 12
Query Match: 63.2% Indels: 34
DB: 2 Gaps: 2

US-10-806-370-11 (1-1748) x US-10-177-661-4 (1-446)
QY 368 ACAACAGTGGGGCTGTACCCATCCATCTCTCCAGGTCAGACCAAGCAACCAGG 427
DB 38 SerProthrinleuSerIleSerSerAspSerSerProIaIaIaIa 57
QY 428 GCCACACAGGAGAGCCAGGTACGAGCTCCCAAGTTCACCTGGCGGAGGGCCAGAG 487
DB 58 LeuLeuLeu-----GlyThrSerLeuProIySerPheThrIrpAsgIuGlnIly 74
QY 488 CAGTACCGGTCTCATCGGGGTGCTGCTCTCTCATTCCTCGGAGGTTTGGCTCATCTC 547
DB 75 GlnLeuProIeuIleGlyCysValLeuLeuLeuIleAlaLeuValIaSerIleIle 94
QY 548 CTCTTCAGTTCGTGACAGGAGCCACAGAGGATCAGGTACAGAGCAGAGAGAGCTGT 607
DB 95 LeuPheGlnPheIrpGlnIlyIleIsthrGlyIleArgIlyIleGlnIlyIleGlnIly 114
QY 608 CCCAAGCAGCCTGTCTGCTGTGACGGGGGTGTGAGTCTGCAAGTGAAGTGAAGCTG 667
DB 115 ProIyHisIleAlaValArgCysAspGlyValIaAspCysIlySerIySerAspGlnLeu 134

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|------|-------------------|----------|----------|---------|------------|---------|---------|---------|----------|---------|---------|-------|-------|-------|-------|--------|--------|-----|------|-----|-----|-----|-----|------|------|------|------|------|------|
| QY | 668 | GGCTGCGGAGAGTTTGA | CTGGGACA | CAAGTCTG | CTGCTAA | AAATCTACTG | GGGCTCC | CCAT | 727 | | | | | | | | | | | | | | | | | | | | | |
| Db | 135 | GLYCSVALA | GRHAE | BPTRP | APRPLYS | SEI | NEU | YSILETY | SEGLYSE | SEHNIS | 154 | | | | | | | | | | | | | | | | | | | |
| QY | 728 | CATGGCTTCC | ATCTGTAG | CAAGCA | CTG | AATGATCT | CTACTA | CGA | GAAGAC | CTGGCCAG | 787 | | | | | | | | | | | | | | | | | | | |
| Db | 155 | GLINTRP | LEUPRO | ILIECY | SERSE | SE | ANTRP | ANAP | SE | TYR | SEGLULV | THCY | 661n | 174 | | | | | | | | | | | | | | | | |
| QY | 788 | CAGCTGGG | TTTGA | AGTGC | TCA | CCCGA | CAAC | CGAG | TTGCC | CA | AGG | GAATTTT | TGCC | AC | 847 | | | | | | | | | | | | | | | |
| Db | 175 | GLIN | EU | GLYP | HEGLU | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 179 | | | | | | | | | | | | | | | |
| QY | 848 | AGCTTCT | CAATCTT | GA | AGAT | CA | ACT | CCAC | CTCA | CGA | AAA | GCCTT | CA | CAG | TTGA | TGC | 907 | | | | | | | | | | | | | |
| Db | 180 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ArgSer | ILucys | 183 | | | | | | | | | | | |
| QY | 908 | CCTTCC | CA | GGG | GTAT | AT | ATCT | CCCTCC | CA | AGT | TTTCC | CA | CTG | CCG | CA | TGA | GGG | CCAT | TGA | CCGG | 967 | | | | | | | | | |
| Db | 184 | PRO | SE | RG | INA | GLY | TRILE | SE | RL | EN | GLIN | Y | SE | SE | HN | ICYS | EG | LY | LEU | ARG | LA | AME | TH | RE | GLY | 203 | | | | |
| QY | 968 | CGA | ATCG | GGG | AGG | GGG | CG | CTG | AG | CT | CG | AT | AG | CA | AG | TGG | CG | CTT | GG | CA | AG | TGA | GTG | AC | 1027 | | | | | |
| Db | 204 | ARG | ILE | VAL | IG | LY | LA | LEU | AL | SE | R | SE | RY | STR | PRO | T | RG | IN | VAL | SE | LEU | HN | IS | 223 | | | | | | |
| QY | 1028 | TTG | GC | CA | CA | CC | CA | AT | CTGT | GA | AG | CA | CG | CT | CA | T | TGA | CG | CC | CA | GTG | GG | GT | CT | CA | TG | CC | 1087 | | |
| Db | 224 | PH | EG | LY | TR | TH | TR | HN | IL | IE | CY | EG | LY | GLY | TH | TR | LEU | LEA | RG | IN | TR | VAL | LEU | TH | LA | 243 | | | | |
| QY | 1088 | GG | CC | CA | CT | CTT | CG | TGA | CC | CG | GG | AGA | AG | CT | CG | GA | GG | GG | CT | GA | AG | GT | TA | CG | GG | CC | 1147 | | | |
| Db | 244 | ALA | HS | CP | HE | PI | NE | VAL | TH | ARG | GLU | YS | VAL | LEU | GLU | GLY | TR | P | LY | VAL | LY | LA | GLY | 263 | | | | | | |
| QY | 1148 | ACC | GA | CA | CA | CT | GA | CA | CA | AG | TTC | CT | GA | GG | CA | CG | CT | CC | AT | TG | CC | GA | T | CA | T | CA | CA | CAGC | 1207 | |
| Db | 264 | TH | SE | R | SE | LEU | HN | IL | EG | IN | LEU | PRO | GLU | LA | LA | SE | RI | LEA | LAG | UL | IE | LA | LE | LA | AN | SE | 283 | | | |
| QY | 1208 | AAT | TA | CA | CC | CA | TGA | GA | GA | GA | GA | GA | CT | AT | TA | CA | CT | CG | CC | CT | CA | TG | CG | CT | CT | CA | AG | CC | CTG | 1267 |
| Db | 284 | AE | NTY | TH | RA | EP | GLU | GLU | ASP | LY | TR | AP | LEA | LA | LEU | ME | TA | ARG | LEU | SE | LY | PRO | LEU | 303 | | | | | | |
| QY | 1268 | ACC | CT | GT | CC | GG | CA | CT | CA | CA | CT | CT | GC | CT | CC | CC | CA | TGA | GA | CA | CA | CT | TTA | AG | CC | CT | 1327 | | | |
| Db | 304 | TH | REU | SE | RA | LA | HS | IL | IE | SP | RO | LA | CY | LEU | PRO | ME | TH | IS | GLY | GLN | TH | PH | SE | RI | LEU | 323 | | | | |
| QY | 1328 | AAT | GA | CA | CT | GT | CG | AT | CA | CA | CG | CT | TTG | CA | GA | CC | CA | GA | GA | CA | GA | CA | TGA | CA | CA | AT | CC | 1387 | | |
| Db | 324 | AE | NG | I | U | TH | RY | ST | TR | PL | LE | TH | GLY | PH | EG | I | Y | ST | TH | RA | RG | I | U | TH | RA | SP | LY | TH | SE | 343 |
| QY | 1388 | CC | CT | TC | CC | GG | GA | GG | GT | CA | AG | T | CA | AT | TC | CA | CTT | CA | GA | AA | TGA | AG | TA | CT | ACT | 1447 | | | | |
| Db | 344 | PRO | PH | EU | LA | RG | I | U | LA | GLN | VAL | AS | NU | LE | AS | PH | EL | YS | CY | AS | | | | | | | | | | |

```

RESULT 4
US-10-177-661-6
; Sequence 6, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Vica, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177,661
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/299,606
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(418)
; OTHER INFORMATION: Xaa = unknown
US-10-177-661-6

Alignment Scores:
Pred. No.: 4,7e-65 Length: 418
Score: 969.50 Matches: 208
Percent Similarity: 56.8% Conservative: 27
Best Local Similarity: 50.2% Mismatches: 130
Query Match: 29.1% Indels: 49
Gaps: 12

US-10-806-370-11 (1-1748) x US-10-177-661-6 (1-418)
QY 602 AGCTGTCACAGACGAGCTGTGCTGTGACGCGGTGTGTGACTGCAGAGTGAAGTATG 661
| | | | | : | | | | | | | | | | | | | | |
DB 1 SerGySlelPro***Ser***TrrpCyAspIylVal***AAspCySPro***GlyGluAsp 20
| | | | | : | | | | | | | | | | | | | | |
QY 662 GAGCTGGGCTGC-----GTGAGGTTTGACTGG 688
| | | | | | | | | | | | | | | | | | | | |
DB 21 Glu*****CysVal*****GlyPro*****ValArgLeuTyrGly 40
| | | | | | | | | | | | | | | | | | | | |
QY 689 GACAAgTCTCGCTTAATATCTACTGCGGTCCTCCATCAGTGCGCTCCCATCTGTAGC 748
| | | | | | | | | | | | | | | | | | | | |
DB 41 Asp*****LeuLeuGlnValTyrSerSerSer*****Trrp**ProValCysSer 60
| | | | | | | | | | | | | | | | | | | | |
QY 749 AGCAACTGATGATGACTCTCTACTCTACAGAGAAGCTGCCAGCAGCTGGATTGGAGAGTCT 808
| | | | | | | | | | | | | | | | | | | | |
DB 61 AspAsnTrrpAnGlnSerTyrSer*****AlaCys***GlnMetGly*****SerIle 80
| | | | | | | | | | | | | | | | | | | | |
QY 809 CACCGAGCAACCGAG-----GTTGCCACAGAGGATTGGCCAAACAGCTTCTCAATCTTG 862
| | | | | | | | | | | | | | | | | | | | |
DB 81 **Tyr***SerGlu***Gly*****Arg***GlyAlaAsnSerPhe***LysLeu 100
| | | | | | | | | | | | | | | | | | | | |
QY 863 AGA-----TCAACTCCACATCCAGGAAGC 889
| | | | | | | | | | | | | | | | | | | | |
DB 101 AsnValSerPro***AsnLeuLeu*****Asp***TyrThrSer**IleGln***** 120
| | | | | | | | | | | | | | | | | | | | |
QY 890 CTCACAGAGGTGTGA-----TGCCCTTCCAGCGGATATATCTCCCTCCAGTGTTC 940
| | | | | | | | | | | | | | | | | | | | |
DB 121 Leu***ArgSerSerSer*****CysProSerGly***ValValSerLeuIndInCysSer 140
| | | | | | | | | | | | | | | | | | | | |
QY 941 -----CACTGCGGACTGAG-----GCCATGACCGCGGCGGATCTGTGGAGGG 982
| | | | | | | | | | | | | | | | | | | | |
DB 141 **GlnAspCysGlyValArgLeuAsnAla*****MetThrSerArgIleValGlyGly 160
| | | | | | | | | | | | | | | | | | | | |
QY 983 GCGCTGCGCTTCGGATAGCAATGCGCTTGGGAAGTGATCTGCACCTTGGGACG---AC 1033
| | | | | | | | | | | | | | | | | | | | |
DB 161 *****AlaSer***Gly***TrrpProTrrpGlnValSerLeuGln*****GlyVal 180
| | | | | | | | | | | | | | | | | | | | |

```

| | | | |
|--|------|--|------|
| QY | 1040 | AACTCTGTGGAGGACGCTCATTTGACGCCAGGGGGGCTCACTGGCGCCACTCGTTC | 1093 |
| | | ::: | |
| Db | 181 | HisIeuCysGlyGlySerLeuIle***Pro***TrrValIleuTrAlaAlaHisCysVal | 200 |
| QY | 1100 | TTCTGTGACCCGGGAGAGTCTCTGGAGGGCTGGAAGGTGTACGGCGGG-----ACGACG | 1153 |
| | | ::: | |
| Db | 201 | ***GlyArg*****LysProIleu***GlyTrp***ValPheAlaGlyIleLeuThr*** | 220 |
| QY | 1154 | AACTGCAACCAAGTTGCTCGAGCGAGCC-----TCCATTGCCAGATCATATCAACAGC | 1207 |
| | | ::: | |
| Db | 221 | SerIeuHis*****Pro***Ala*****Arg***ValGlyLysIleIleIleHisPro | 240 |
| QY | 1208 | AATTAC-----ACGATGAGGAGACGCACTGTGACATCGCCCTCATG | 1249 |
| | | ::: | |
| Db | 241 | AsnTrp*****Ser*****LysAspAsnAspIleAlaIleuLeu | 260 |
| QY | 1250 | CGAGCTGTCCAGACCCCTGACCCCTGCGCTCACATATCCACCTGCTTGCTGCCATGCAT | 1309 |
| | | ::: | |
| Db | 261 | LysIeuSerLysProLeuThrPhe***AspTrpIleGlnProValCysIleuProAsnPro | 280 |
| QY | 1310 | GGACAGACCTTTAAGCTTCATATGACACCTGCTGATCACAGGCTTTGGCAACACAGGAG | 1368 |
| | | ::: | |
| Db | 281 | GlyIleGln***Leu***ProGlyThrThrCysTrpIle***GlyTrpGlyAlaThr***Glu | 300 |
| QY | 1370 | ACAGATGACAAAGACATCCCCCTCTCTCCGGAGGTGACAGTGCATCTCATGACTTCAAG | 1429 |
| | | ::: | |
| Db | 301 | *****GlyTrpTrpSerProValIleuGlnIleAla***ValProIleuIleAspAsnLys | 320 |
| QY | 1430 | AAATGCAATGACTCTTGCTATGACAGTTAACCTTACCCCAAGATGATGTCGTGG | 1489 |
| | | ::: | |
| Db | 321 | ***CysAsnSerTrp***ValTrpAspAsn***IleIleProAlaGlyMetIleCysAlaGly | 340 |
| QY | 1490 | GACCTGTGGGGGCGAGACACTCTCTCCAGGAGACAGCGGGGGCTTGTCTGTGAG | 1549 |
| | | ::: | |
| Db | 341 | TyrIleuGlnGlyGlyValAspSerCysGlnGlnIleAspSerGlyGlyProIleuValCysIleu | 360 |
| QY | 1550 | -----CAGAACACCGCGCTGTAACCTGGCAGGTGTCCACCACTGGGGGCGACAGGC | 1597 |
| | | ::: | |
| Db | 361 | *****GlnAsnAsnAlaArgTrpIleu***Gly***ThrSerTrpGly***Gly | 380 |
| QY | 1598 | TGT---GGCCAGAGAAACAAACCTGTGTGTGTACCAAGTGACAGAGATTCTTCCCTGG | 1654 |
| | | ::: | |
| Db | 381 | Cys***AlaIleValAlaSerLysProGlyValIleTrp***ValIle***PheIleu***Trp | 400 |
| QY | 1655 | ATTTACAGCAAGATGAGAGCGAGGTGGATTCAGAAAATCC | 1696 |
| | | ::: | |
| Db | 401 | IleTyrSerGlnMet***AlaGln***ArgPheArgLysSer | 414 |
| RESULT 5 | | | |
| : Sequence 932, Application US/09759143 | | | |
| : Patent No. 6800746 | | | |
| : GENERAL INFORMATION: | | | |
| : APPLICANT: Xu, Jiangchun | | | |
| : APPLICANT: Dillon, Devin C. | | | |
| : APPLICANT: Mitcham, Jennifer L. | | | |
| : APPLICANT: Harlocker, Susan L. | | | |
| : APPLICANT: Jiang, Yugu | | | |
| : APPLICANT: Henderson, Robert A. | | | |
| : APPLICANT: Kalos, Michael D. | | | |
| : APPLICANT: Fanger, Gary R. | | | |
| : APPLICANT: Retter, Marc W. | | | |
| : APPLICANT: Stolk, John A. | | | |
| : APPLICANT: Day, Craig H. | | | |
| : APPLICANT: Vedvick, Thomas S. | | | |
| : APPLICANT: Carter, Derrick | | | |
| : APPLICANT: Li, Samuel | | | |
| : APPLICANT: Wang, Aijun | | | |
| : APPLICANT: Skeiky, Yasir A.W. | | | |
| : APPLICANT: Hepler, William | | | |
| : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND | | | |
| : FILE REFERENCE: 210121.427C23 | | | |

```

1 CURRENT APPLICATION NUMBER: US/09/759,143
2 CURRENT FILING DATE: 2001-01-12
3 NUMBER OF SEQ ID NOS: 934
4 SOFTWARE: FastSeq for Windows Version 3.0
5 SEQ ID NO 932
6 LENGTH: 492
7 TYPE: PRT
8 ORGANISM: Homo sapiens
9 US-09-759-143-932
10
11 Alignment Scores:
12 Pred. No.: 6,47e-50 Length: 492
13 Score: 769.50 Matches: 178
14 Percent Similarity: 50.6% Conservative: 74
15 Best Local Similarity: 35.7% Mismatches: 195
16 Query Match: 23.1% Indels: 51
17 Db: 2 Gaps: 15
18
19 US-10-806-370-11 (1-1746) x US-09-759-143-932 (1-492)
20
21 QY 288 GGTCTTCATCGCGCAGGTGAT--CATCCGCGAGGTACAGCTCGGTGACAACTCCCA 344
22 |||||::: |||||::: |||||::: |||||::: |||||:::
23 Db 12 G1PProTyrTYrGIuaenH1sg1TYrGIuPProGIuaenPro-----TyPProAlaGIu 29
24
25 QY 345 CCNAAGGT-----ACCTGTGTAGACAAACACAGTGGGGGCGTAAACCATCC----- 392
26 |||||::: |||||::: |||||::: |||||::: |||||:::
27 Db 30 ProThrValValProThrValTYrGIuValH1sProAlaG1nTYrProSerProVal 49
28
29 QY 393 GATCATCTCCGCGAGGTCCAGACCAACCAACAGGCGCACAGGAGAAAGCCAGGTACGA 452
30 |||||::: |||||::: |||||::: |||||::: |||||:::
31 Db 50 ProG1nTYrAlProAlaGValLeuThrG1nAlaSerAnPro-----ValValCys 66
32
33 QY 453 GCCTGCCCAAGTTACCTGGCGGGAGGCCAAGACGTACCGCTATCGGTCGTCG 512
34 |||||::: |||||::: |||||::: |||||::: |||||:::
35 Db 67 ThrG1nProLysSerProSerGIYThrValCYthrSerLysThrLysAlaLeuCYs 86
36
37 QY 513 -----TCCTCCTCATTCGCCCTGG 530
38 |||||::: |||||::: |||||::: |||||::: |||||:::
39 Db 87 I1eThrLeuThrLeuGIYThrPheLeuValGIYAla1aLeuAla1aGIYLeuLeuTYr 106
40
41 QY 531 TGGTTTGCTCATCATCTCTTCAGATTCTGGCGGCGCACACAGGATCAGTACGAAG 590
42 |||||::: |||||::: |||||::: |||||::: |||||:::
43 Db 107 LysPheMetGIYSerLysCYsSerAnSerGIY1aGIuCYsAbpSerSerGIYThr--- 125
44
45 QY 591 AGCAGAGGAGAGAGCTGTCCCAAGCACGCTGTTCGTGTGACGGGTGGTGACTGCAAC 650
46 |||||::: |||||::: |||||::: |||||::: |||||:::
47 Db 126 -----Cys11eAnPProSerAnThrP--CysAbpGIYValSerH1sCYsProG 141
48
49 QY 651 TGAAGAGTACAGAGCTGGCGTCGTGAGGATTTCAGTGGGACAAAGTCTCTGTTAAATCT 710
50 |||||::: |||||::: |||||::: |||||::: |||||:::
51 Db 141 IYGIYGIuAspGIuAsnArGIYsValArGIYLeuTYrGIYSerAnPhe11eLeuAlnValT 161
52
53 QY 711 ACTCTGGGCTCTCCCATCATGAGGTGCTCCCATCTGTAGACAGCACTGGAATATCTCTACT 770
54 |||||::: |||||::: |||||::: |||||::: |||||:::
55 Db 161 YrSerSerGIuArGIYsSerTYrP1sProValCYsGIuAspSerTYrPsm1c1uAnTYrG 181
56
57 QY 771 CAGAGAAGACCTCGCACAGCTGGGTTCGAGATGTCTCACCGGACAAACGAGGTTGCC 830
58 |||||::: |||||::: |||||::: |||||::: |||||:::
59 Db 181 IYArGIAla1aCYArGIYsAbpMerCIYrLysAnAnPheTYrSerSerC1n---GIYI 200
60
61 QY 831 ACAGGAGATTGGCCAAACAGCTTCAATCTTGAGATACAACTCCACC----- 877
62 |||||::: |||||::: |||||::: |||||::: |||||:::
63 Db 200 LeVal1aAspSerGIYsSerThrSerPheMetLysLeuAnThSer1aGIYAnValA 220
64
65 QY 878 --ATCCAGGAAGAGCTCCACAGGTGAA--TGCCCTTCCACAGCGGATATATCCCTCC 932
66 |||||::: |||||::: |||||::: |||||::: |||||:::
67 Db 220 sP1eTYrLYsLYsLeuTYrH1sSerAnPAlaCYsSerSerGIYAlaValValSerLeuA 240
68
69 QY 933 AGTGTCTCCCACTGGCGACTGAGG-----GCCATGACCGGGCGGATCTGGGAGGGG 983
70 |||||::: |||||::: |||||::: |||||::: |||||:::
71 Db 240 rGcYs11eAlaCYsGIYValAsnLeuAnSerSerArGIuInSerArGIY1eValGIYGIY 260
72
73 QY 984 CGCTGGCCTCGGATAGCAAGTGGCCTTGGCAAGTGAAGTCTGACTTGGCACCAACCA 1043

```

```

Db      260  luseRAlaleuProGlyAlaTriProTriGlnValSerLeuHisValGlnAsnValHisv 280
Qy      1044  TCTGTGAGGACGACCTCATTTGACGCCCGGAGGCTTCACATGCGCCACGACTCTTCTCG 1103
Db      280  aICySgLyglYserIleIleThrProGluTriIleValThnAlaHisCysVal----- 298
Qy      1104  TGACCCGGAGAGAGTCTCTGAGAGG-----TGAAAGTGTACGGCGGACCGACGA 1154
Db      299  -----GluTyrProLeuAsnAsnProTriPhtsTriPhtAlaPheIleGlyIleLeuA 316
Qy      1155  ACCTGCACCACTTCCCTGAGGACGCC-----TCCATTGCCGAATCATCTCAACAGCA 1208
Db      316  rglInserPheMetPheTyrGlyAlaGlyTyrGlnValGluTyrValIleSerHisProA 336
Qy      1209  ATTACACCGATGAGAGAGACATGACATGACATGCGGCTGATCGGCTGTCACCGCCCTGA 1268
Db      336  enTyrAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
Qy      1269  CCTGTCCGCTCATCATCCACCTGCTGCTGCCATGATGATGACAGACCTTTAGCCTCA 1328
Db      356  hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnProG 376
Qy      1329  ATGAGACCTGCTGATCATCAGGCTTTGGCAAAGACCGAGAGACAGATGACAGCATCC 1388
Db      376  luginLeuCystrPrlSerGlyTyrGlyAlaThrGluGlu---LysGlyLysThrSerG 395
Qy      1389  CCTTCCTCCGGAGAGTGAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTGCTGG 1448
Db      395  luValIleAsnAlaAlaLysValLeuLeuIleGluThrGlnArgCysAsnSerArgTyrV 415
Qy      1449  TCTATGACAGTACTTAACCCCAAGATGATGTGTGTGGGACCTTGTGGGGGACAG 1508
Db      415  aITyrAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnIleYsnValA 435
Qy      1509  ACTCTCTCCAGAGGACAGCGGGGGGCTCTTGTCTGTGAGCAAAACAACCGCTGTGTACC 1568
Db      435  spSerCySgGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTriPrl 455
Qy      1569  TGGAGAGTGTACAGCTGGGGGACAGGCTGTGTGTGTGAGCAAAACAACCGCTGTGTACA 1628
Db      455  euIleGlyAspThrSerTriPrlSerGlyCysAlaLysAlaTyrArgProGlyValTyrG 475
Qy      1629  CCAAGAGTACAGAAATTCTTCCCTGATTTTACAGCAAGATGAGAGCGAG 1678
Db      475  lyaEnValMetValPheThrAspTriPrlIeTyrArgGlnMetArgAlaAsp 491

```

```

APPLICANT: Fanger, Gary R.
APPLICANT: Wantanabe, Yoshihiro
APPLICANT: Mesgher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-932

Alignment Scores:
Pred. NO.: 6,47e-50 Length: 492
Score: 769.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
DB: 2 Gaps: 15

US-10-806-370-11 (1-1748) x US-10-012-896-932 (1-492)
Qy      288  GGTCTCATCCGCGAGGTCAT---CATCCCGCAGTCAAGCTCGGTGACAACTCCCAA 344
Db      12  GLyProTyrTyrGluAsnHisGlyTyrGlnProGluAsnPro-----TyrProAlaGln 29
Qy      345  CCAGAGTGT-----ACCTGTTAGACAAACCCAGTGGGGGCTGTACCATCC----- 392
Db      30  ProThrValValProThrValTyrGluValHisProAlaGlnTyrTyrProSerProVal 49
Qy      393  GATCATCTCCTGCAGGTCAAGACCAAGCAACGAGGCGCACAGGAGAGACCCAGGTACGA 452
Db      50  ProGlnTyrAlaProAlaGValLeuThrGlnAlaSerAsnPro-----ValValCys 66
Qy      453  GCTGCCCCAAGTTCACCTGCGGGAGGCGCAAGACAGTACCGCTCATCGGTCGCTGC 512
Db      67  ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysAlaLeuCys 86
Qy      513  -----TCCTCTCATTTGCCCTGG 530
Db      87  IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaGlyLeuLeuTyr 106
Qy      531  TGGTTTCGCTCATCATCTCTCCAGTTCGTGGGAGGCGCACACAGGATCAGGTACAG 590
Db      107  LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125
Qy      591  AGCAGAGGAGAGCTGTCCCAAGACACCGCTTGCCTGTGACGAGGTGTGAGTGCACAGC 650
Db      126  -----CysIleAsnProSerAsnTyr---CysAspGlyValSerHisCysProG 141
Qy      651  TGAAGAGTACAGAGCTGGCTGCTGAGTTTGTGACTGGAGACAGTCTGTCTTAAATCT 710
Db      141  LysGlyGluAspGluAsnArgCysValArgLeuTyrGlySerAsnPheIleLeuGlnValT 161
Qy      711  ACTGTGCTCTCCATCAGTGGCTTCCCATCTGTGATGACAGATGATGATGATCTTACT 770
Db      161  TyrSerSerGlnArgLysSerTriPhtsProValCysGlnAspAspTriPhtsGlnAsnTyrG 181
Qy      771  CAGAGAACCTGCGACAGCTGGGTTTGCAGAGTGTCCACCGGACACCGAGGTTGCC 830
Db      181  LysArgAlaAlaCysArgAspMetGlyTyrLysAsnAsnAsnPheTyrSerSerGln---GlyI 200
Qy      831  ACAGGATTTTCCCAACAGCTTCTCAATCTTGAGATCAACTCCACC----- 877
Db      200  LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 220
Qy      878  --ATCCAGGAAGGCTCAGAGGTGGA---TGCCCTTCCCAAGGCGGTATATGCTCCCTCC 932
Db      220  spIleTyrLysLysLeuTyrHisSerAspAlaCysSerSerLysValValValSerLeuA 240

```

| | | | |
|----|------|---|------|
| Qy | 933 | AGTGTCCCACTGGGGAGTAAAG-----GCATGACCGGGCGAGTGTGGGAAGGG | 983 |
| | | | |
| Db | 240 | rgCysIIleAlaCySGlYValAsnLeuAsnSerSerArgInserArgIleAlaIdylGlyc | 260 |
| Qy | 984 | CGCTGGCCCTCGGAATGACAAAGTGGCTTTGGCAAGTAGCTGTGCACCTTGGGCAACCCACA | 1043 |
| | | | |
| Db | 260 | IuSerAlaIleuProGlyAlaIleThrProIleGlnValSerLeuHisValGlnAsnValHisV | 280 |
| Qy | 1044 | TCGTGTGAGGCAAGCTCATTTGAAGCCCAAGTGGGTCTCATCTGGCCGCCCATCTTTCTTCG | 1103 |
| | | | |
| Db | 280 | alCySGlYglYSerIleIleThrProGluTrpIleValThrAlaAlaHisCySValJ---- | 298 |
| Qy | 1104 | TGACCCCGGGAAGAAGTCTGTGAAGGGC-----TGGAAGGTGATGACGGGCAACAGCA | 1154 |
| | | | |
| Db | 299 | -----GluIlyProLeuAsnAsnProTrpHisIleThrAlaIlePheAlaIdylIleuA | 316 |
| Qy | 1155 | ACCTGCACACAGTTGCTCTGAGGACAGCC-----TCCATTGGCCGAATCATCTCAACAGCA | 1208 |
| | | | |
| Db | 316 | rgGlnSerPheMetPheTrpIolYAlaGlyTrgInValGlnIlyValIleSerHisProA | 336 |
| Qy | 1209 | ATTACACCGATGAGAGGAGGACGACTATGATGACATGCGCCCTCATGCGCTGTCCAAAGCCCTGA | 1268 |
| | | | |
| Db | 336 | snTrYAspSerIyThrIySAsnAsnSprIleAlaIleMetIyLeuGlnIlySProLeuT | 356 |
| Qy | 1269 | CCCTGTCCGCTCAATCCACCCCTGTCCTGCCCCCAATGACATGACAGACACTTGAAGCTCA | 1328 |
| | | | |
| Db | 356 | hrPheAsnAspLeuValIlySProValCysIleProAsnProGlyIleMetLeuGlnProG | 376 |
| Qy | 1329 | ATGAGACCTGTGTGATCAGAGCTTTGGCAAGACAGGAGAGACAGATGACAAACATCC | 1388 |
| | | | |
| Db | 376 | IuGlnIleuCyStrIpleSerIlyTrpGlyAlaThrGlnGlu--LysIlyIyStrHisArg | 395 |
| Qy | 1389 | CCTTCTCCGGGAGGTGCAGGTCAATCTCATGCACTTCAAGAAATGCATGACTTACCTGG | 1448 |
| | | | |
| Db | 395 | IuValIleuAsnAlaAlaIySValIleuIleuIleGluThrGlnArgCysAsnSerArgTyv | 415 |
| Qy | 1449 | TCTATGACAGTACTTACCTTACCCCAAGAGATGTGTGCTGGGGGACCTTGTGTGGGGCAGAG | 1508 |
| | | | |
| Db | 415 | alTrYAspAsnIleuIleThrProAlaMetIleCySAlaGlyPheIleuGlnIlyAsnValA | 435 |
| Qy | 1509 | ACTCTCTCCAGAGGAGACAGCGGGGGGCTCTTGTCTGTGACAGAAACAACCGCTGTAC | 1568 |
| | | | |
| Db | 435 | spSerCySGlnIlyAspSerIlyGlyProIleuValThrSerIySAsnAsnIleTrpTrpL | 455 |
| Qy | 1569 | TGGCAGGTGTCCACCAAGCTGGGGCACAGGCTGTGGCCAGAGAAACAACCTGTGTGTACA | 1628 |
| | | | |
| Db | 455 | eulIleGlyAspThrSerTrpIySerIyGlySAlaIySAlaIyArgProGlyIyAlTyrg | 475 |
| Qy | 1629 | CCAAAGTGACAAAGTCTTCCCTCGAGATTACAGCAAGATGGAAGCGAG | 1678 |
| | | | |
| Db | 475 | IyAsnValMetValPheThrAspTrpIleTyArgGlnMetArgIleAsp | 491 |

```

RESULT 7
US-10-144-678A-932
: Sequence 932. Application US/10144678A
: Patent No. 7033827
:
: GENERAL INFORMATION:
:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yugu
: APPLICANT: Henderson, Robert A.
: APPLICANT: Kalos, Michael D.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Reiter, Marc W.
: APPLICANT: Stolk, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aijun
:

```

| | | | |
|----|--|--|------------------|
| | APPLICANT: | Skeiky, Yasir A. W. | |
| | APPLICANT: | Hepner, William T. | |
| | APPLICANT: | Hural, John | |
| | APPLICANT: | McNeill, Patricia D. | |
| | APPLICANT: | Houghton, Raymond L. | |
| | APPLICANT: | Vinals y de Basols, Carlota | |
| | APPLICANT: | Foy, Teresa M. | |
| | APPLICANT: | Matanabe, Yoshihiro | |
| | APPLICANT: | Deryn, Ta | |
| | TITLE OF INVENTION: | COMPOSITIONS AND METHODS FOR THE THERAPY AND | |
| | TITLE OF INVENTION: | DIAGNOSIS OF PROSTATE CANCER | |
| | FILE REFERENCE: | 210121.427C28 | |
| | CURRENT APPLICATION NUMBER: | US/0/144,678A | |
| | CURRENT FILING DATE: | 2002-08-12 | |
| | NUMBER OF SEQ ID NOS: | 1033 | |
| | SOFTWARE: | FastSeq for Windows Version 3.0 | |
| | SEQ ID NO 932 | | |
| | LENGTH: | 492 | |
| | TYPE: | PRT | |
| | ORGANISM: | Homo sapiens | |
| | US-10-144-678A-932 | | |
| | Alignment Scores: | | |
| | Pred. No.: | 6.47e-50 | Length: 492 |
| | Score: | 769.50 | Matches: 178 |
| | Percent Similarity: | 50.6% | Conservative: 74 |
| | Best Local Similarity: | 35.7% | Mismatches: 195 |
| | Query Match: | 23.1% | Indels: 51 |
| | DB: | 3 | Gaps: 15 |
| QY | US-10-806-370-11 (1-1748) x US-10-144-678A-932 (1-492) | | |
| QY | 288 GGTCTTCATCCGGGACGATCAT--CATCCGCCAGGTCAAGCTCGTGACAACCTCCCA 344 | | |
| | ::: :: | | |
| Db | 12 GLyProTyrTYrGLuAsnHISGLyTrGlnProGluAsnPro-----TyProlaGln 29 | | |
| QY | 345 CCAGAAGT-----ACCTTTTAAGACAACACCAGTGGGGGCTGAACCATTC----- 392 | | |
| | ::: :: | | |
| Db | 30 ProThrValValProThrValTyrGluValHisProAlaGlnTyrTyrProSerProval 49 | | |
| QY | 393 GATCATCTCCGCAGGTCAGACACAGCACAGGGGCCACCGAGGAGACCCAGGTACGA 452 | | |
| | ::: :: | | |
| Db | 50 ProGlnTyrAlaProIArgValLeuThrglnAlaSerAsnPro-----ValValCys 66 | | |
| QY | 453 GCCTGCCCAAGTTTCACCTGCGCGGAGGGCCAGAACAGCATCCGCTCATCGGGTGGTC 512 | | |
| | ::: :: | | |
| Db | 67 ThrGlnProIySerProSerOlyThrValCysThrSerLethrLysAlaLeuCys 86 | | |
| QY | 513 -----TCTCTTCATTGCCCTGG 530 | | |
| | ::: :: | | |
| Db | 87 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaIleGlyLeuLeuTyr 106 | | |
| QY | 531 TGGTTTGGCTATCATCTCTTCCAGTGTGCGACAGGCCACACAGGAGTAGTAGAG 590 | | |
| | ::: :: | | |
| Db | 107 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125 | | |
| QY | 591 AGCAGAGGAGAGTGTCTCCCAAGCACAGCTGTTCGCTGTGATCGGGGTGTGACTGCAAGC 650 | | |
| | ::: :: | | |
| Db | 126 -----CyAlaIleAsnProSerAsnTyrP--CysAspGlyValSerHisCysProG 141 | | |
| QY | 651 TGAAAGTAGCAGAGCTGGGGCTGCGTAGAGGTTTGACTGGGACAAAGTCTCTGCTTAAATCT 710 | | |
| | ::: :: | | |
| Db | 141 LygLygluAspGluAsnArgCysValArgTyrGlySerAsnPheIleLeuGlnValIt 161 | | |
| QY | 711 ACTCGGGGTCTCCCATCAGTAGGCTTCCCATCTGTGACAGCAACCTGGAATGACTCTACT 770 | | |
| | ::: :: | | |
| Db | 161 TyrSerSerGlnArgCysArgSerTyrPHisProValCysGlnAspHisAspTyrPanGluAsnTyrG 181 | | |
| QY | 771 CAGAGAAACCTGACAGCAGAGCTGGGTTTTGAGAGTGCTTACCCGACCAACCGAGTTGCC 830 | | |
| | ::: :: | | |
| Db | 181 LyArgAlaAlaCysArgArgAspMetGlyTyrTyrLysAsnAsnPheTyrSerSerGln--GlyI 200 | | |
| QY | 831 ACAGGAGATTTGGCCAACAGCTTCTCAATCTTGAAGATACACTTCAC----- 877 | | |

Db 260 luseralaleuProglYalatrProtrpoglValserleuHleValGlnasnValHisv 280
QY 1044 TCCTGGAGGAGCAGCTTATGAGCCCAAGGGGTCTCATGCCGCCACTGCTTCTTCG 1103
Db 280 alCysglYglYserlelleThrProglutPrleValThrAlaAlaHisCysVal---- 298
QY 1104 TGACCCGGGAGAGGCTCGTAGGGC-----TGGAGGTCTAGCGGGCAGCAGCA 1154
Db 299 -----GluYsProleuasnAnProtrpHisTrpThrAlaPheAlaGlylleuA 316
QY 1155 ACCTGCACAGTTCCTGAGGAGCC-----TCCATTGCCGAGATCATCATCAAGCA 1208
Db 316 rgGlnserPheMetPheTyrGlyAlaGlyTyrGlnValGlnYValIleSerHisProA 336
QY 1209 ATTAACACCATGAGAGAGAGCATATGACATCGCCCTCATAGCGCTGTCCAGGCCCTGA 1268
Db 336 snYrAsPserLysThrLysasnAsnApriLeAlaMetLysLeuGlnLysProLeut 356
QY 1269 CCGGTGCGGTGCATCCACCCCTGCTTGGCTCCCATGACATGACAGACCTTAGCTCA 1328
Db 356 hrPheasnAAspLeuValLysProValCysLeuProasnProglYMetMetLeuGlnProG 376
QY 1329 ATGAGACCTGCTGATCATCAGGCTTTGGCAAGACAGAGAGACAGATGACAGACATCCC 1388
Db 376 luginleuCystrpIleSerGlyTrpGlyAlaThrGlnGlu---LysGlyLysThrserG 395
QY 1389 CTTTCTCCGGAGAGTGCAGGTCAATCTCATGACTTCAAGAAATGCAATGACTACTTGG 1448
Db 395 luValleuAsnAlaAlaLysValLeuLeuIleGlnThrGlnArgCysAsnSerArgTyrv 415
QY 1449 TCATATGACATTACCTTACCCCAAGATGATGTGTGTGGGACCTTCGTGGGGCAGAG 1508
Db 415 alYrAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 435
QY 1509 ACTCTCCAGGAGAGAGAGCGGGGCTCTTGTCTGTAGCAGAGAAACCGCTGTACC 1568
Db 435 epserCySglnGlyAspSerGlyGlyProleuValThrserLysAsnAsnIleTrpIPL 455
QY 1569 TGGCAGGTGTACAGCTGGGGCAGAGGCTGTGGCCAGAGAAACCTGTGTGTACA 1628
Db 455 eulieGlyAspThrserTrpGlySerGlyCysAlaLysAlaTyrArgProglYalTyrG 475
QY 1629 CCAAAGGACAGAGTCTTCCCTGGATTTACACCAAGTCGAGAGCGAG 1678
Db 475 lYasnValMetValPheThrAspTrpIleTyrArgGlnMetArgAlaAsp 491

RESULT 9
US-09-691-840-2
; Sequence 2, Application US/09691840
; Patent No. 644419
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavligian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TPRS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/691,840
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US/09/342,749
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/091,044
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-691-840-2

Alignment Scores: 7.7e-50 Length: 492

Score: 768.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
DB: 2 Gaps: 15

US-10-806-370-11 (1-1748) x US-09-691-840-2 (1-492)

QY 288 GGTTCATATCCGGCAGGTCTAT---CATCCGCAGTCAAGCTCGGTACAACTTCCCGCA 344
Db 12 GlyProTyrTrpGlnAsnHisGlyTyrGlnProGlnAsnPro-----TyrProAlaGln 29
QY 345 CCAAGAGT-----ACCTGTAGAGCAACACAGTGGGGGTGTACCATCC----- 392
Db 30 ProThrValValProThrValTyrGlnValHisProAlaGlnTyrTrpProserProval 49
QY 393 GATCATCTCTGCGCAGTGCAGACCAAGCAACAGGGCCACAGGAGAGCCAGCA 452
Db 50 ProGlnTyrAlaProArgValLeuThrGlnAlaSerAsnPro-----ValValCys 66
QY 453 GCGTCCCAAGTTCACCTGGCGGGAGGCGCAAGACAGTACCGCTATCGGTGCGTCC 512
Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysValAlaLeuCys 86
QY 513 ----- 530
Db 87 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp 106
QY 531 TGGTTTGGCTCATATCATCTCTCTCCAGTTCTGAGAGGCGCACAGGATGAGTACAG 590
Db 107 LysPheMetCylserLysCysSerAsnSerGlyLleGluCysAspSerSerGlyThr--- 125
QY 591 AGCAGAGGAGAGCTGCCCAAGACAGCTTGTGCTGTGACGGGTGTGACTGCAGC 650
Db 126 -----CysIleAsnProSerAsnTrp---CysAspGlyValSerHisSerProG 141
QY 651 TGAAGAGTGCAGAGCTGGCTGGGTGAGTTTGTCTGGACAGTCTGTGTTAAATCT 710
Db 141 LysGlyLysAspGlyAsnValArgValArgLeuTyrGlyProAsnPheIleLeuGlnValT 161
QY 711 ACTCTGGGTCCTCCATCATGAGCTTCCCATCTGTAGACAGCAACTGTAGTACTCTACT 770
Db 161 YrserSerGlnArgLysSerTrpHiserProValCysGlnAspAspTrpAsnGlnAsnTyrG 181
QY 771 CAGAGAAAGCTGCAGACAGCTGGGTTTGCAGAGTCTCACCGGACACAGGAGTTGCC 830
Db 181 lYArgAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
QY 831 ACAGGATTTTGGCAACAGCTTCTCAATCTTGATGATCACTCCACC----- 877
Db 200 leValAspAspSerGlySerThrPheMetLysLeuAsnThrSerAlaGlyAsnValA 220
QY 878 --ATCCAGAAAGCTTCACAGGCTTGA---TGCCCTCCACAGCGGTATATCTCCCTCC 932
Db 220 spIleTyrLysLysLysLeuTyrHiserThrAspAlaCysSerSerLysAlaValValSerleuA 240
QY 933 AGTGTTCACCTGGGAGCTAGC-----GCCATGACCGGGCGGATGTGGAGGGG 983
Db 240 rGcysIleAlaCysGlyValAlaAsnLeuAsnSerSerArgLysAspArgIleValGlyG 260
QY 984 CGCTGGCTGGCATAGCAAGTGTGCAAGTGTGCAAGTGTGCACTTGCGGACCAACCCACA 1043
Db 260 luseralaleuProglYalatrProtrpoglValserleuHleValGlnasnValHisv 280
QY 1044 TCCTGGAGGAGCAGCTTATGAGCCCAAGGGGTCTCATGCCGCCACTGCTTCTTCG 1103
Db 280 alCysglYglYserlelleThrProglutPrleValThrAlaAlaHisCysVal---- 298
QY 1104 TGACCCGGGAGAGGCTCTGAGAGC-----TGGAGGTCTAGCGGGCAGCAGCA 1154
Db 299 -----GluYsProleuasnAnProtrpHisTrpThrAlaPheAlaGlylleuA 316
QY 1155 ACCTGCACAGTTCCTGAGGAGCC-----TCCATTGCCGAGATCATCATCAAGCA 1208

Db 316 rginserPhemePheTyrglYalaglYrYrglnValGluLysValIleSerHisPro 336
1209 ATTACCCGATGAGAGAGAGATGACATGACGCGCCGTCATGCGGTCGCAAGCCCGGA 1268
Db 336 snTyraPserLysThrLysAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
1269 CCTGTCGCGATCACATCCACCCCTGCTGCTCCCATGATGACAGACCTTTAGCCTCA 1328
Db 356 hrPheAsnAspLeuValLysProValCyLeuProAsnProGlyMetLeuGlnProG 376
1329 ATGAGACCTGCTGATCATCAGGCTTTGGCAAGACGAGGAGACGATGACAGACATCCC 1388
Db 376 LnglnLeuCyStrpIleSerGlyTrpGlyAlaThrGluGln---LysGlyLysThrSerG 395
1389 CCTTCCTCCGGAGAGGTGATGATCATGACCTTCAAGAAATGCAATGACTACTGG 1448
Db 395 LuValLeuAsnAlaAlaLysValLeuLeuIleGlnThrGlnArgCyAsnSerArgTyrv 415
1449 TCTATGACAGTTACCTTACCCCAAGATGATGTGTGCGGACCTTGTGCGGAGCAAG 1508
Db 415 alTyraPAsnLeuIleThrProAlaMetIleCyAlaGlyPheLeuGlnGlyAsnValA 435
1509 ACTCTGCGCAAGGAGACAGCGCGCGCTTGTGTGTGAGACAAACAACCGCTGTACC 1568
Db 435 spSerCyGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTrpTrpL 455
1569 TGGCAGGTGTACAGAGGTGGGAGCAGGCTGTGGCAGAAACAAACCTGTGTGTACA 1628
Db 455 euIleGlyAspThrSerTrpLysSerGlyCyAlaLysValLysTyraArgProGlyValTyrg 475
1629 CCAAGTACAGAAAGTTCTTCTCGATTTTACAGCAAGATGAGAGCGAG 1678
Db 475 LysAsnValMetValPheThrAspTrpIleTyraArgGlnMetArgAlaAsp 491

RESULT 10

US-09-949-016-11074
Sequence 11074, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 11074
LENGTH: 510
TYPE: PRT
ORGANISM: Human
US-09-949-016-11074
Alignment Scores:
Pred. No.: 7.8e-50 Length: 510
Score: 768.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
Gaps: 15

US-10-806-370-11 (1-1748) x US-09-949-016-11074 (1-510)

QY 288 GGTCTCATCGGAGGAGTAT---CATCGCAGGTGACGCTGGTGAACACTCCCA 344
Db 30 GlyProTyrrTyrgLysAsnHisGlyTrpGlnProGlnAsnPro-----TyrrProAlaGln 47

QY 345 CCAGAGTGT-----ACCTTGTAGACCAACACCAAGTGGGGCTGTACCCATCC----- 392
Db 48 ProThrValValProThrValTyrgLysValHisProAlaGlnTyrrProSerProVal 67
QY 393 GATATCTCTCCGAGGTCAAGACCAACCAAGGCGCCACGAGGAAGACCCAGTACA 452
Db 68 ProGlnTyrrAlaProAspGValLeuThrGlnAlaSerAsnPro-----ValValCyS 84
QY 453 GCGTCCCAAGTTACCTGGCGGAGGCGCAAGACGACTACCGCTGACGCGGTGCGTGC 512
Db 85 ThrGlnProLysSerProSerGlyThrValCyThrSerLysThrLysValLeuCyS 104
QY 513 -----TCCTCTCAATTCCTGG 530
Db 105 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaIleGlyLeuLeuTrp 124
QY 531 TGGTTTGCCTCATGATCTCTTCCAGTTCTGGACGGGCCACACAGGATCAGTACAAAG 590
Db 125 LysPheMetGlySerLysCySserAsnSerGlyIleGluCyAspSerSerGlyThr--- 143
QY 591 AGCAGAGGAGAGGCTGCCAAGACCGCTGTGCTGTCGACGGGCTGTGACTGCAAGC 650
Db 144 -----CySLeuAsnProSerAsnTrp---CySAspGlyValSerHisCySProG 159
QY 651 TGAAGATGACAGACCTGGCTGCGTGAAGTTTGACTGGACAACTCTGCTTAAATCT 710
Db 159 LysGlyLysAspGlnAsnArgCySValArgLeuTyrgLysProAsnPheIleLeuGlnValT 179
QY 711 ACTTGAGTCTCTCCATCAAGTGTCTCCCATCTGTACAGCAACACTGGATGACTCTACT 770
Db 179 YrSerSerGlnArgLysSerTrpHisProValCySglnAspAspTrpAsnGlnAsnTyrg 199
QY 771 CAGAGAACCTGCGCACACCTGGGTTTGCAGAGTGTCTCACCGGACCAACGAGGTTGCC 830
Db 199 LysArgAlaAlaCySArgAspMetGlyTyrrLysAsnAsnPheTyrrSerSerGln---GlyI 218
QY 831 ACAGAGATTTGCCCAACAGCTTTCATCTTGAGATCAACTCCACC----- 877
Db 218 LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 238
QY 878 --ATCCAGAAAGCTCCACAGCTTGA---TGCCTTCCCAAGCGGTATATCTCCCTCC 932
Db 238 spIleTyrrLysLeuLysIleThrHisSerAspAlaCySserSerLysValAlaValSerLeuA 258
QY 933 AGTGTCCCACTGGCGGACTAGG-----GCCATGACCGCGCGGATCGTGGAGAGG 983
Db 258 rgCySleAlaAlaCySgLyValAsnLeuAsnSerSerArgGlnSerArgIleValGlyGlyG 278
QY 984 CGCTGCGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTGTGCACTTGGCACCACCA 1043
Db 278 LysSerAlaLeuProGlyAlaTrpProTrpGlnValSerLeuHisValGlnAsnValHisV 298
QY 1044 TCTGTGAGGACAGCTATGAGCGCCAGTGGGCTGCTGACGCGGCCCATGCTTCTTGG 1103
Db 298 alCySgLyGlySerLleIleThrProGlnTrpIleValThrAlaAlaHisLeuValA--- 316
QY 1104 TGACCCGGAGAGAGTCTCTGAGAGGC-----TGAAGTGTATACCGGCGACCA 1154
Db 317 -----GluLysProLeuAsnAsnProTrpHisTrpThrAlaPheAlaGlyIleLeuA 334
QY 1155 ACCTGACCAAGTTCCTGAGGACGC-----TCCATTCGCGAGATCATCATCAACGA 1208
Db 334 rginserPhemePheTyrgLyalaglYrYrglnValGluLysValIleSerHisProA 354
QY 1209 ATTACCCGATGAGAGAGAGATGACATTCGCGCTCAACGCGCTGTCCAGCCCTGA 1268
Db 354 snTyraPserLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 374
QY 1269 CCTGTCGCGTCAATCCACCTGCTGCTCCCATGATGACGATGACAGACCTTTAGCCTCA 1328
Db 374 hrPheAsnAspLeuValLysProValCyLeuProAsnProGlyMetLeuGlnProG 394

Db 415 a1TyrAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 435
QY 1509 ACTCTGTCACAGAGACAGACGGGGGCTTGTCTGTGAGACAGACACCGCTGGTACC 1568
Db 435 epSerCysGlnGlyAspSerGlyGlyProLeuValThrSerAsnAsnLeuIleTrpL 455
QY 1569 TGGCAGGTGTACCAAGCTGGGGGACAGGCTGTGGCCAGAAACAACTGTGTGTACA 1628
Db 455 euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaLysArgProGlyValTyrG 475
QY 1629 CCAAGCAGACAGAGTCTTCCCTGGATTTACAGCAAGATGAGAGC 1675
Db 475 LysAsnValMetValPheThrAspTrpIleTyrArgGlnMetLysAla 490

RESULT 14
US-09-759-143-895
Sequence 895, Application US/09759143
Patent No. 6800746
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasar A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
NUMBER OF SEQ ID NOS: 934
SOFTWARE: PaedSeq for Windows Version 3.0
SEQ ID NO 895
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-895

Alignment Scores:
Pred. No.: 1,55e-49 Length: 492
Score: 764.50 Matches: 178
Percent Similarity: 50.5% Conservative: 73
Best Local Similarity: 35.8% Mismatches: 195
Query Match: 22.9% Indels: 51
Gaps: 15

US-10-806-370-11 (1-1748) x US-09-759-143-895 (1-492)

QY 288 GGTCTCATCGGAGAGTCAAT---CATCCGCCAGAGTCAAGCTCGGTGACAACTCCCAA 344
Db 12 G1ProTyrTyrGluAsnHisGlyTyrGlnProGluAsnPro-----TyrProAlaGln 29
QY 345 CCAGAGTGT-----ACCTTGTTAGAGCAACACAGTGGGGGCTGTACCCATCC----- 392
Db 30 ProThrValValProThrValTyrGluValHisProAlaGlnTyrTyrProSerProVal 49
QY 393 GATCATCTCTCCAGAGTCAACAGACAGCAACAGAGGACAGAGGAGAGCCAGGTACGA 452
Db 50 ProGlnTyrAlaProArgValLeuThrGlnAlaSerAsnPro-----ValLysCys 66
QY 453 GCTGCGCCAAATTCACTGGCGGGAGGGCCAGAAAGCAAGCTTACCGCTCATCGGCTGC 512

Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerIleThrLysLysAlaLeuCys 86
QY 513 -----TCCTCTCATATGGCCCTGG 530
Db 87 IleThrIleuThrIleuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp 106
QY 531 TGGTTTGCTCATATCCCTCTTCCAGTCTGGACGGGCCACACAGAGGATCAGGTACAG 590
Db 107 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125
QY 591 AGCAGAGAGAGAGCTGTCCCAAGACACCTGTTCCTGTGACGGGGTGTGACCTGCACAG 650
Db 126 -----CysIleAsnProSerAsnTrp---CysAspGlyValSerHisCysArg 141
QY 651 TGAAGAGTACAGAGCTGGGCTGCGTGGAGTTTACTGGACCAAGTCTGTCTTAATCT 710
Db 141 LysGlyGluAspGluAsnArgCysValArgLeuTyrGlyProAsnPheIleLeuGlnMet 161
QY 711 ACTCTGGGCTCTCCATCAAGTGGCTTCCCATCTGTAGACAGCACTGGAAATGACTTACT 770
Db 161 YSerSerGlnArgLysSerTrpHisProValCysGlnAspAspTrpAsnGluAsnTyrG 181
QY 771 CAGAGAGAGCTGCGCAGAGCTGGGTTTCAGAGAGTCTCACCGAGCAACGAGGTTCCC 830
Db 181 LysArgAlaAlaCysArgAspMetLysTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
QY 831 ACAGGATTTGGCAACAGCTTCCATCTTGAATATCACTCCACC----- 877
Db 200 LevalAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 220
QY 878 --ATCCAGAAAGCTCCACAGGTTGAA--TGCCCTTCCACCGGTATATCTCCCTCC 932
Db 220 spIleTyrLysLysLeuTyrHisSerAspAlaCysSerSerLysAlaValAlaSerLeuA 240
QY 933 AGTGTCCCACTGGAGCTAGG-----GCCATGACCGGGCGGATGTGGAGAGGG 983
Db 240 TgCysLeuAlaCysGlyValAlaAsnLeuAsnSerSerArgLysSerArgLysGlyG 260
QY 984 CGCTGGCCTCGATAGCAATGGGCTTGGAGAGTGTGACTGTGACCTGGACCCACCA 1043
Db 260 LysSerAlaLeuProGlyAlaIleTrpTrpTrpAlaValSerLeuHisValGlnAsnValHis 280
QY 1044 TCTGTGGAGGACGCTCATTTGACGCCAGTGGTGTCTACTGCCGCCACTGCTTCTTG 1103
Db 280 alCysGlyGlySerIleIleThrProGluTrpIleValThrAlaAlaHisCysVal---- 298
QY 1104 TGACCCGGAGAGAGTCTTGAAGGC-----TGAAGGTGTAGCGGGACCCAGCA 1154
Db 299 -----GluLysProLeuAsnAsnProTrpHisTrpThrAlaPheAlaGlyIleLeuA 316
QY 1155 ACCTGACACAGTTGCTGAGAGCAGCC-----TCCATTTGCCAGATCATCATCAAGCA 1208
Db 316 rGlnInsrrPheMetPheTyrGlyAlaGlyTyrGlnValGlnLysValIleSerHisProA 336
QY 1209 ATTAACACCGATAGAGAGAGCACTATGACATCGCCCTCATCGGCTGTCCAAAGCCCTGA 1268
Db 336 snTyrAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeu 356
QY 1269 CCTGTCCGCTCATCATCAACCTGCTGCTTCCCTCCCATGATGACAGACAGCACTTAAGCTCA 1328
Db 356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnProG 376
QY 1329 ATGAGACCTGTGATCAACAGGCTTGGCAAGACAGAGGAGAGACAGATGACAGCATCC 1388
Db 376 LysGlnLeuCysTrpIleSerGlyTyrGlyAlaThrGlnGlu---LysGlyLysThrSerG 395
QY 1389 CCTTCTCCGAGAGTGCAGTCAATCTCATGCAATGCAAGAAATGCAATGACTACTTG 1448
Db 395 LuValLeuAsnAlaAlaLysValLeuLeuIleGluThrGlnArgCysAsnSerArgTyr 415
QY 1449 TCTATGACAGTATCACTTACCCCAAGATGATGTGTGTGGGACCTTGTGGGGGAGAG 1508
Db 415 alTyrAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 435

QY 1509 ACTCTGCGAGGAGACAGAGGGGGCTCTTGTCTGTAGACAGACCGCTGTACC 1568
 Db 435 spherCysgInglYAspSerGlyGlyProLeuValThrSerAsnAsnAsnIleTrrIprL 455
 QY 1569 TGGCAGGTGTCCACGAGCTGGGGGACAGAGCTGTGGCCAGAGAAACAACTGTGTGTACA 1628
 Db 455 eullegIAspThrSerTrrpGlySerGlyCysAlaLysAlaTyrArgProGlyValTyrG 475
 QY 1629 CCAAGTGACAGAAAGTTCTTCCCTGTGATTTCACAGCAAGATGAGAGC 1675
 Db 475 LysnValMetValPheThrAspTrrIleTyrArgInMetLysAla 490
 RESULT 15
 US-10-012-896-895
 / Sequence 895, Application US/10012896
 / Patent No. 6943236
 / GENERAL INFORMATION:
 / APPLICANT: Xu, Jiangchun
 / APPLICANT: Dillon, Devin C.
 / APPLICANT: Mitcham, Jennifer L.
 / APPLICANT: Harlocker, Susan L.
 / APPLICANT: Jiang, Yugu
 / APPLICANT: Kalos, Michael D.
 / APPLICANT: Retter, Marc W.
 / APPLICANT: Stolk, John A.
 / APPLICANT: Day, Craig H.
 / APPLICANT: Vedvick, Thomas S.
 / APPLICANT: Carter, Darick
 / APPLICANT: Li, Samuel X.
 / APPLICANT: Wang, Aijun
 / APPLICANT: Skeiky, Yasar A.W.
 / APPLICANT: Hepler, William T.
 / APPLICANT: Henderson, Robert A.
 / APPLICANT: Hural, John
 / APPLICANT: McNeill, Patricia D.
 / APPLICANT: Houghton, Raymond L.
 / APPLICANT: Vinals de Basols, Carlota
 / APPLICANT: Foy, Teresa
 / APPLICANT: Fanger, Gary R.
 / APPLICANT: Mantanabe, Yoshihiro
 / APPLICANT: Meagher, Madeleine Joy
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 / FILE REFERENCE: 210121.427C27
 / CURRENT APPLICATION NUMBER: US/10/012.896
 / CURRENT FILING DATE: 2001-12-10
 / NUMBER OF SEQ ID NOS: 1011
 / SOFTWARE: FASTSEQ for Windows Version 3.0
 / SEQ ID NO 895
 / LENGTH: 492
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-012-896-895
 Alignment Scores:
 Pred. No.: 1.55e-49 Length: 492
 Score: 764.50 Matches: 178
 Percent Similarity: 50.5% Conservative: 73
 Best Local Similarity: 35.8% Mismatches: 195
 Query Match: 22.9% Indels: 51
 DB: Gaps: 15
 US-10-806-370-11 (1-1748) x US-10-012-896-895 (1-492)
 QY 288 GGTCTATCGGAGGAGTCAAT---CATCCGCGAGGTACGCTCGGTGACCACTCCCA 344
 Db 12 GilyProtylYrGluAsnHisGlyTyrGlnProGluAsnPro-----TyrProAlaGln 29
 QY 345 CCAGAGTGT-----ACCTGTGTAGACCAACCAAGTGGGGCTGTACCCATCC----- 392
 Db 30 ProthrValValProthrValTyrGluValHisProAlaGlnTyrTyrProSerProval 49

QY 393 GATCATCTCTGCGAGGTGACAGCAACGAGGCCACGAGGAGAGCCAGTAGCA 452
 Db 50 ProGlnTyrAlaProAlaValLeuThrGlnAlaSerAsnPro-----ValAlaCys 66
 QY 453 GCCTGCCCAAGTTCACCTGCGGGGAGGAGCCAGAAAGCAAGCTACCGCTCATGGGTGCTGC 512
 Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysAlaLeuCys 86
 QY 513 -----TCCTGCTCATTTGCCCTGG 530
 Db 87 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaIleuAlaLeuAlaGlyLeuLeuTrr 106
 QY 531 TGGTTTCGCTCATCATCTCTTCCAGTTCTGCGAGGCCACACAGGAGTACGATACAG 590
 Db 107 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125
 QY 591 AGCAGAGGAGAGCTGTCCCAACAGCAGCTTGTGCTGTGACGGGGGTGTGACTGCACG 650
 Db 126 -----CysIleAsnProSerAsnTrr-CysAspGlyValSerHisCysProG 141
 QY 651 TGAAGAGTACAGAGCTGGCTGGGTGAGTTGACTGGACAGAGTCTCTGCTTAATCT 710
 Db 141 LylGlyGluAspGluAsnArgCysValArgLeuTyrGlyProAsnPheIleLeuGlnMet 161
 QY 711 ACTGTGGTCTCTCCATCATGAGCTTCCATCTGTAGACAGCAACTGGAATGACTCTACT 770
 Db 161 YrsSerSerGlnArgLysSerTrrPheProValCysGlnAspAspTrrAsnGlnAsnTyrG 181
 QY 771 CAGAGAAAGCTGCGACAGCAGCTGGGTTTCGAGAGTGTCAACCGGACACCGAGTTGCC 830
 Db 181 LysArgAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
 QY 831 ACAGGATTTTGGCAACAGCTTCTCAATCTTGAGATCAACTCCACC----- 877
 Db 200 LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerValGlyLysnValA 220
 QY 878 --ATCCAGAAAGCTCCACAGGCTGTGA---TGCCCTTCCACGCGGTATATCCCTCC 932
 Db 220 splleTrrLysLysLeuTyrHisSerAspAlaCysSerSerLysAlaValValSerLeuA 240
 QY 933 AGTGTTCACATCGGAGCTGAGG-----GCCATGACCGGGCGGATGTGGAGGGG 983
 Db 240 rGysLeuAlaCysGlyValAsnLeuAsnSerSerArgInSerArgIleValGlyIlg 260
 QY 984 CGCTGCGCTCGATAGCAAGTGGCTTGGCAAGTGAATCTGCACCTTGGCACCCACACA 1043
 Db 260 luserAlaLeuProGlyAlaTrrProTrrPheGlnValSerLeuHisValGlnAsnValHisv 280
 QY 1044 TCTGTGAGGACAGCTCATTTGAGCGCCAGTGGGAGTGCACATGCCGCCACTGTCTTGG 1103
 Db 280 alCysGlyGlySerIleIleThrProGlnTrrIleValIlnAlaAlaHisCysVal---- 298
 QY 1104 TGACCCGAGAAAGTCTGAGAGGC-----TGAAAGTGTATACCGGGCACACAGA 1154
 Db 299 -----GluLysProLeuAsnAsnAsnProTrrPheIleTrrPheAlaPheAlaGlyIleLeuA 316
 QY 1155 ACCTGACCAAGTGTCTGAGCAGCC-----TCCATTTCCGAGATCATATCAACAGCA 1208
 Db 316 rGlnSerPheMetPheTyrGlyAlaGlyTyrGlnValGlnLysValIleSerHisProA 336
 QY 1209 ATTTACACCATGAGAGAGAGCAATGACATCGCCCTTCATGCGGCTGTCCAGGCCCTGA 1268
 Db 336 snTyrAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
 QY 1269 CCTGTCCGCTCATCATCCACTGTCTTGCCTCCCATGATGACAGACCTTTAGCCTCA 1328
 Db 356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnProG 376
 QY 1329 ATGAGACCTGCTGTGATACAGGCTTTGGCAAGACAGCAGGAGACAGATGACAAAGCATCC 1388
 Db 376 lueGlnLeuCystrIleSerGlyTrrpGlyAlaThrThrGlnGlu---LysGlyLysThrSerG 395
 QY 1389 CCTTCTCTCGGAGAGTGCAGGTCAATCTCATTCGACTTCAAGAAATGCAATGACTACTTG 1448

Db 395 luValleuAsnAlaAlaLysValleuLeuileGluThnGlnArgCyAsnSerArgTyr 415

QY 1449 TCTTGTGACAGTACTTACCCCAAGAGATGATGTGTGCGGGAACCTTGTCGGGCGACAG 1508

Db 415 aATyAspAsnLeuileThnProAlaMetileCyAsnAGlyPheLeuIndmIyAsnValA 435

QY 1509 ACTCTGTGCAGAGAGACAGCGGGGGCTCTTGTCTGTGACAGACAACCGCTGTGACC 1568

Db 435 spSerCyGlnGlyAspSerIyGlyProLeuValThnSerAsnAsnIleTriprPL 455

QY 1569 TGGAGAGGTGCACAGCTGTGGGCGACAGCGTGTGGCAGAAACAAACCTGTGTGTACA 1628

Db 455 euileGlyAspThnSerIprGlySerIyGlySaluIySaluIyArgProIyValIyArg 475

QY 1629 CCAAGTGCACGAAGTCTTCCCTGGATTTTCACAGACATGGAAGC 1675

Db 475 lyaAsnValMetValPheThnAspTriIeIyArgGlnMetIyValA 490

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 02:28:48 ; Search time 58.1 Seconds
(without alignments)
4180.893 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 3333
Sequence: 1 cccagagaccatcgagagag99.....ggctctgtgactcgagaa 1748

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2097797 seqs, 463214858 residues
Total number of hits satisfying chosen parameters: 4195594

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1p
-O=/abs/ABSSWB.spool/US10806370/runat.15092006.105646.10654/app_query.fasta_1
-DB=Published Applications AA Main -OEMT=fastan -SUFFIX=n2p.rapbm
-MIMATCH=0.1 -LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTMT=ptco -NORM=ext
-HEAPSIZE=500 -MTLEN=0 -MAXLEN=2000000000 -HOST=abs802p
-USER=US10806370 @CCN 1.1 446 @runat.15092006.105646.10654 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:
1: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US10_PUBCOMB.pep:*
5: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------------------------------|
| 1 | 3004 | 90.1 | 569 | 4 | US-10-428-275-130 Sequence 130, App |
| 2 | 2999 | 90.0 | 562 | 3 | US-09-879-792-12 Sequence 12, App1 |
| 3 | 2999 | 90.0 | 562 | 4 | US-10-806-370-12 Sequence 12, App1 |
| 4 | 2991 | 89.7 | 562 | 4 | US-10-156-214A-24 Sequence 24, App1 |
| 5 | 2991 | 89.7 | 562 | 5 | US-10-916-758-94 Sequence 94, App1 |
| 6 | 2963 | 88.9 | 688 | 4 | US-10-156-214A-26 Sequence 26, App1 |
| 7 | 2963 | 88.9 | 688 | 5 | US-10-916-758-2 Sequence 2, App1 |
| 8 | 2956 | 88.7 | 581 | 4 | US-10-353-690-100 Sequence 100, App |
| 9 | 2953.5 | 88.5 | 586 | 4 | US-10-428-275-120 Sequence 120, App |
| 10 | 2948.5 | 88.5 | 586 | 4 | US-10-428-275-104 Sequence 104, App |
| 11 | 2948.5 | 88.5 | 586 | 4 | US-10-428-275-124 Sequence 124, App |

| | | | | | |
|----|--------|------|-----|---|-------------------------------------|
| 12 | 2943.5 | 88.3 | 586 | 4 | US-10-428-275-142 Sequence 142, App |
| 13 | 2942.5 | 88.3 | 586 | 4 | US-10-428-275-140 Sequence 140, App |
| 14 | 2775.5 | 83.3 | 537 | 3 | US-09-888-615-104 Sequence 104, App |
| 15 | 2775.5 | 83.3 | 537 | 4 | US-10-428-275-128 Sequence 128, App |
| 16 | 2775.5 | 83.3 | 537 | 6 | US-11-037-243-104 Sequence 104, App |
| 17 | 2691 | 80.7 | 542 | 4 | US-10-428-275-122 Sequence 122, App |
| 18 | 2291.5 | 68.8 | 477 | 4 | US-10-177-661-2 Sequence 2, App1 |
| 19 | 2291.5 | 68.8 | 477 | 5 | US-10-910-507-2 Sequence 2, App1 |
| 20 | 2287 | 68.6 | 421 | 3 | US-09-898-837A-15 Sequence 15, App |
| 21 | 2287 | 68.6 | 421 | 4 | US-10-428-275-132 Sequence 132, App |
| 22 | 2219.5 | 66.6 | 466 | 4 | US-10-274-639-8 Sequence 8, App1 |
| 23 | 2219.5 | 66.6 | 466 | 4 | US-10-333-574-8 Sequence 8, App1 |
| 24 | 2130.5 | 63.9 | 442 | 4 | US-10-428-275-144 Sequence 144, App |
| 25 | 2108 | 63.2 | 446 | 4 | US-10-177-661-4 Sequence 4, App1 |
| 26 | 2108 | 63.2 | 446 | 5 | US-10-910-507-4 Sequence 4, App1 |
| 27 | 2097 | 62.9 | 382 | 4 | US-10-428-275-134 Sequence 134, App |
| 28 | 2059 | 61.8 | 406 | 4 | US-10-428-275-126 Sequence 126, App |
| 29 | 2059 | 61.8 | 406 | 4 | US-10-428-275-138 Sequence 138, App |
| 30 | 2026 | 60.8 | 401 | 4 | US-10-428-275-112 Sequence 112, App |
| 31 | 2016 | 60.5 | 401 | 4 | US-10-428-275-116 Sequence 116, App |
| 32 | 2012 | 60.4 | 401 | 4 | US-10-428-275-118 Sequence 118, App |
| 33 | 1842.5 | 55.3 | 370 | 4 | US-10-428-275-114 Sequence 114, App |
| 34 | 1839.5 | 55.2 | 367 | 4 | US-10-428-275-110 Sequence 110, App |
| 35 | 1474 | 44.2 | 283 | 4 | US-10-428-275-148 Sequence 148, App |
| 36 | 1311 | 39.3 | 296 | 3 | US-09-804-156-30 Sequence 30, App1 |
| 37 | 1311 | 39.3 | 296 | 4 | US-10-067-761-30 Sequence 30, App1 |
| 38 | 1311 | 39.3 | 296 | 4 | US-10-319-519-30 Sequence 30, App1 |
| 39 | 1311 | 39.3 | 372 | 3 | US-09-804-156-15 Sequence 15, App1 |
| 40 | 1311 | 39.3 | 372 | 3 | US-09-946-633-7 Sequence 7, App1 |
| 41 | 1311 | 39.3 | 372 | 4 | US-10-125-459-7 Sequence 7, App1 |
| 42 | 1311 | 39.3 | 372 | 4 | US-10-067-761-15 Sequence 15, App1 |
| 43 | 1311 | 39.3 | 372 | 4 | US-10-319-519-15 Sequence 15, App1 |
| 44 | 1270 | 38.1 | 239 | 4 | US-10-428-275-106 Sequence 106, App |
| 45 | 1270 | 38.1 | 239 | 4 | US-10-428-275-108 Sequence 108, App |

ALIGNMENTS

RESULT 1
US-10-428-275-130
; Sequence 130, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OR INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; SOFTWARE: Cureseq1st version 0.1
; SEQ ID NO 130
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-275-130
Alignment Scores:

Pred. No.: 8.36-175 Length: 569
Score: 3004.00 Matches: 563
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.1% Indels: 0
DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x US-10-428-275-130 (1-569)

```
QY 8 ACCATGGAGAGGACAGCCACGGGAATGATCTCCAGCAAGAACCTTCAGTGGAGCA 67
DB 4 ThiretGUAAGAASerHisAGLYAsnAlaSerProAlaArgThrProSerAlaGLYAla 23
QY 68 TCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTCCAGGCCGGGATCTCCAGCCAGCA 127
DB 24 SerProAlaGlnAlaSerProAlaGLYThrProProGlyArgAlaSerProAlaGlnAla 43
QY 128 TCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTCCGGGCCGGGATCTCCAGCCAGCA 187
DB 44 SerProAlaGlnAlaSerProAlaGLYThrProProGlyArgAlaSerProAlaGlnAla 63
QY 188 TCTCCAGCTGGTACCTCCAGGCCGGGATCTCCAGGCCGGGATCTCCAGCCAGCA 247
DB 64 SerProAlaGlnAlaSerProAlaGLYThrProProGlyArgAlaSerProAlaGlnAla 83
QY 248 TCTCCAGCCCGAGGATCTCCAGCTGGAGACACCTTCAGGATCTCCAGCCAGCA 307
DB 84 SerProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerSerGlyArgSer 103
QY 308 TCATCCCGCAGGTCAGCTCGGTGACAACTCCCAACAGAGTACCTTGTATAGCA 367
DB 104 SerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAla 123
QY 368 ACACGAGTGGGGGTGACCATCCATCATCTCTGACAGTACAGCAACCAAG 427
DB 124 ThrProAlaGlnAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArg 143
QY 428 GCCACAGGAGAGCCACAGTACAGACCTGCCAAGTTACCTGACGGAGAGGACAGAG 487
DB 144 AlaThrArgGlnSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlyGlnLys 163
QY 488 CAGCTACCGCTCATCGGGTGGTGGTCTCTCATTCGCTGGTGGTGGTCTCATCATC 547
DB 164 GlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValaSerLeuIleIle 183
QY 548 CTCTTCCAGTCTGGCAGGAGCAGCAGAGATCAAGTCAAGAGAGGAGAGAGCTG 607
DB 184 LeuPheGlnPheTrpGlnGlyHisThrGlyIleArgGlyLysGlnArgGlnSerCys 203
QY 608 CCCAAGCAGCTGTTGCTGCTGACCGGGTGGTGAAGTGAAGTGAAGTGAAGCTG 667
DB 204 ProLysHisAlaAlaValArgCysAspGlyValaValaAspCysLysLeuLysSerAspGlnLeu 223
QY 668 GCGTGGCTGAGGTTGATCGGACAAAGTCTGCTTAAATCTACTCTGGTCTCCAT 727
DB 224 GlnCysValaArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHis 243
QY 728 CAGTGGCTTCCATCTGTAAGAGCAACTGGAATGACTCTTAAGAGAGAAAGCTGGCAG 787
DB 244 GlnTrpLeuProIleCysSerSerSerAsnTrpAsnAspSerTyrSerGlnLysThrCysGln 263
QY 788 CAGCTGGTTCAGAGTGTACACCGGACAAACGAGTGGCCACAGGAGATTTGGCCAAC 847
DB 264 GlnLeuGlyPheGlnSerAlaHisArgThrArgValaAlaHisArgAspPheHisAsn 283
QY 848 AGCTTTCATCTTGAGATACAACTCCACATCCAGGAAAGCTCCACAGGTTGAATGC 907
DB 284 SerPheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuHisArgSerGlnCys 303
QY 908 CTTCCCAAGGGATATCTCCCTCCAGTGTCCCACTGGCGAGCTGAAGGGCATACCGGG 967
DB 304 ProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGly 323
```

```
QY 968 CGATGCTGGAGAGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCAC 1027
DB 324 ArgIleValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValaSerLeuHis 343
QY 1028 TTGGGACCAACCCACATCTGTGAGAGGACCGCTCATTAAGCCCAAGTGGGGTCACTGCC 1087
DB 344 PheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValaLeuThrAla 363
QY 1088 GCCCATGCTTCTTCGTGACCCGGGAGAAAGTCTGAGAGGCTGGAAAGTGTACGCGGCG 1147
DB 364 AlaHisCysPhePheValaThrArgGlnLysValaLeuGlnGlyTrpLysValaTyrAlaGly 383
QY 1148 ACCAGCAACTGACCAAGTGGCTGAGGACAGCTTCATTCGCGAGATCATCAAGAC 1207
DB 384 ThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleLeuAsnSer 403
QY 1208 AATTACACGATAGAGAGAGCACTATAGATGCGCTCATGGGCTGTCCAGCCGCTG 1267
DB 404 AsnTyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeu 423
QY 1268 ACCGTGCTCGCTCATCAACCTGCTTGGCTCCCATGATGAGACAGACCTTGAAGCTC 1327
DB 424 ThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeu 443
QY 1328 AATGACACTGCTGGATCAAGGCTTTGGCAAGCCAGGAGACAGATGACAAACATCC 1387
DB 444 AsnGlnThrCysTrpIleThrGlnGlyPheGlyLysThrArgGlnTrpAspLysThrSer 463
QY 1388 CCGTCTCGGGAGGAGTGCAGTCAATCTGATGACTTCAAGAAATGCAATGATACCTTG 1447
DB 464 ProPheLeuArgGlnValaGlnValaAsnLeuIleAspPheLysCysAsnAspTyrLeu 483
QY 1448 GTCTATGACAGTTCATCTTACCCCAAGATGATGTGTCTGGGAGACTTGTGGGGAGCA 1507
DB 484 ValTyrAspSerTyrLysLeuThrProArgMetCysAlaGlyAspLeuArgGlyArg 503
QY 1508 GACTCTGCCAGGAGACAGCGGGGGCTCTTGTCTGTGACGCAAGAACACCGCTGGTAC 1567
DB 504 AspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyr 523
QY 1568 CTGGCAGTGTGACACCTGAGGAGCAGAGGCTGTGGCCAGAAACAAACCTGTGTATAC 1627
DB 524 LeuAlaGlyValThrSerTrpGlyThrGlyCysGlyAlaTrpGlnLysProGlyValTyr 543
QY 1628 ACCAAGTGAAGAAAGTCTTCCCTGAGATTATACAGAAATGAGAGAGAGTGCATTC 1687
DB 544 ThrIleValThrGlnValaLeuProTrpIleTyrSerLysMetGlnSerGlnValaArgPhe 563
QY 1688 AGAAATCC 1696
DB 564 ArgLysSer 566
```

RESULT 2
US-09-879-792-12
Sequence 12, Application US/09879792
Patent No. US20020061850A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
APPLICANT: Gedrich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Sertine
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879, 792
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211, 224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283, 353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283, 648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT
PRIOR FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 36

(Docket No. US20020061850A1 LIO-81-WO)

; SOFTWARE: FastSEQ for Windows Version 4.0

LENGTH: 562

ORGANISM: Homo sapiens
US-09-879-792-12

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 1.67e-174 | Length: | 552 |
| Score: | 2999.00 | Matches: | 5630 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Best Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 90.0% | Indels: | 0 |
| DB: | 3 | Gaps: | 0 |

US-10-806-370-11 (1-1748) X US-09-879-792-12 (1-562)

| | | | |
|----|-----|--|-----|
| QY | 11 | ATGAGAGAGGACAGCCACCGGAATGATCTCCAGCAAGAAACACCTTCAGCTGGAGCATCT | 70 |
| Db | 1 | MetGluAArgSerSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer | 20 |
| QY | 71 | CCAGCCACGAGCATCTTCAGCTGGGACACCTCCAGGCCGGGACATCTCCAGCCAGGCATCT | 130 |
| Db | 21 | ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer | 40 |
| QY | 131 | CCAGCCACGAGCATCTTCAGCTGGGACACCTCCGGGGCCGGGACATCTCCAGCCAGGCATCT | 190 |
| Db | 41 | ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer | 60 |
| QY | 191 | CCAGCTGGTACACCTTCAGCGCGGGGATCTCCAGGCCGGGACATCTCCAGCCAGGCATCT | 250 |
| Db | 61 | ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer | 80 |
| QY | 251 | CCAGCCCGGGGACATCTCCGGGCTCTGGGATCATCTTCAGGTCTCATCTCCGAGGTCATCA | 310 |
| Db | 81 | ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer | 100 |
| QY | 311 | TCCGCGAGGTCAAGCTTCGGGTGACAACTCCCAACGAGATGTAACCTGTGTAAAGCAACA | 370 |
| Db | 101 | SerAlaArgSerAlaSerValThrThrSerProThrArgAlaThrLeuValAlaArgAlaThr | 120 |
| QY | 371 | CCAGTGGGGGCGTATACCATCCGATCATCTCCGCCAGGTCAAGCACCAGCAACAGGGCC | 430 |
| Db | 121 | ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla | 140 |
| QY | 431 | ACCAGGGAGAGCCCAAGTACGAGACCTGCCCAAGTTAACTCTGGCGGAGGCCAAGACAG | 490 |
| Db | 141 | ThrArgGlnSerProGlyThrSerLeuProGlyPheThrTrpArgGlnGlyGlnGln | 160 |
| QY | 491 | CTAACGGCTATAGGGGAGGTGCTCCCTTCATTTGCCCTGGTTCGCTCATCATCTTC | 550 |
| Db | 161 | LeuProLeuIleGlyCysValLeuLeuIleAlaLeuValaValSerLeuIleIleLeu | 180 |
| QY | 551 | TTCCAGTTTGGCAGGGCCACACAGGATCAGTACAGAGAGCAGAGGAGAGCTGTCCC | 610 |
| Db | 181 | PheGlnPheTrpGlnGlyHisThrGlyIleArgTyrIleSulGlnIleArgGlnSerCysPro | 200 |
| QY | 611 | AAGCAGCTGTTCCTGCTGTGAACGGGGGTGTGTGAATCTCAAGCTGAAGATGACAGCTGGGC | 670 |
| Db | 201 | IysHisAlaIalaValArgCysAspGlyValaIlaAspCysIleLeuIysSerIleAspGlnLeuGly | 220 |
| QY | 671 | TGCGTGAAGTTGACTGGGACAGTCTGTGCTTAATAATCACTCTGGGCTCTCCCATCAG | 730 |
| Db | 221 | CysValArgPheAspTrpAspIlySerLeuLeuIleAlaIleTyrSerGlySerSerHisGln | 240 |
| QY | 731 | TGGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCTACTCAGAGAAAGACTGTGCACAG | 790 |
| Db | 241 | TrpLeuProIleCysSerSerIleAsnTrpAsnAspSerTyrSerGlyIlyThrCysGlnGln | 260 |
| QY | 791 | CTGGCTTTGAGAGTGTCTCACCGGACAAACCGAGTTGCCACAGGATTTTGGCAACAGC | 850 |
| Db | 261 | LeuGlyPheGlnSerIleHisArgGlnThrGlnValAlaIleHisArgAspPheAlaIleSer | 280 |

| | | | |
|--|------|---|------|
| QY | 851 | TTCTCAATCTTTGAGATACAACTCCACATCCAGGAAAGCCCTCCACAGTCTGAAATGCCCT | 910 |
| Db | 281 | PhserIleuuarIgrIyAsmSerThrIleGlnIuSerIeuIshAysSerGluCySPro | 300 |
| QY | 911 | TCCACAGCGATATATCTCCCTCCAGTGTCTCCACATCGCGAGCTGAAGGCCATGACCGGCGG | 970 |
| Db | 301 | SerGlnnrgIyTrIleSerIeuGlnCySerHieCyseIyIeuArGlaIamethnGlyArg | 320 |
| QY | 971 | ATCGTGGAGGGGCGCTGGCCCTCGGATAGCAAGTGGCCTTTGGCAAGTGAAGTCTTGACATTC | 1030 |
| Db | 321 | IleValIglyIalAlaIeuAlaSerAspSerIyStrProItrpGlnValaSerIeuHieSph | 340 |
| QY | 1031 | GGCAACACCCCAATCTGTGGAGGACCGCTCATTTGACGCCCAAGTGGGSGTCCACGCGCGCC | 1090 |
| Db | 341 | GlyThrThrHsiIleCySgIyGlyThrIeuIleAspAlaGlnItrpValIleThrAlaIa | 360 |
| QY | 1091 | CACGCTCTCTTCGTGACCCCGGGAAGAGTCTCGAGAGGGCTGGAAAGTACGCGGCGACCC | 1150 |
| Db | 361 | HieCySphIeuIeuAlThrArgGluIyValIeuGlnIyItrpIySvalTyraIaGlyThr | 380 |
| QY | 1151 | AGCAACCTGCACCAAGTGGCTTGAGGACAGCCCTCCATTGCGGAGATCATATCAACAGCAAT | 1210 |
| Db | 381 | SerIeuHieIsgIuIeuProGluAlaIaSerIleIaGlnIleIleIeAsnSerAsn | 400 |
| QY | 1211 | TACACCGATGAGAGAGACACTATATGACATCGCCCTCATTCGGCTGTCCAGCCCTGACC | 1270 |
| Db | 401 | TyrThrAspGluGluAspAspTyraSpirIleAlaIeuMetAglIeuSerIySProIeuThr | 420 |
| QY | 1271 | CTGTCCGCTCATATCAACCTCGTTCGCTCCCAATGACATGGACATGGACACTTATAGCTCAAT | 1330 |
| Db | 421 | IeuSerAlaHsiIleHsiProAlaCySleuProMetHsiGlyGlnThrPhseIleuAsn | 440 |
| QY | 1331 | GAGACCTGCTGATATCAACAGGCTTTGGCAAGACCAAGGAGACAGATGACAAGACATCCCC | 1390 |
| Db | 441 | GluThrCyStrpIleThrGlnIyPheGlyIySerThrArgIuThrAspAspIyStrSerPro | 460 |
| QY | 1391 | TTCTCCGGGAGGTGACAGGTCAATCTCATGACTTTCAAAGAAATGCATAGACTTCTTGTC | 1450 |
| Db | 461 | PheIeuArGluValIaGlnValaIasnIeuIleAspPheIyIyScyAsnAspTyraIeuVal | 480 |
| QY | 1451 | TATAGACGTTATCTTATACCCCAAGATGATGTGTCCTGGGAGCCTTCGCGGGGCGAGAGAC | 1510 |
| Db | 481 | TyrAspSerTyraIeuThrProArMetMetCysAlaGlyAspIeuArGlyIyArgAsp | 500 |
| QY | 1511 | TCCAGCCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGACAAACCGCTGTACCTG | 1570 |
| Db | 501 | SerCySglnGlyAspSerGlyGlyProIeuValaIySgIuGlnAsnAsnArgItrpTyraIeu | 520 |
| QY | 1571 | GCAGGTGTACACAGCTGGGGGACACAGGCTGTGGCCAGAGAAACAACCTGGTGTATCAC | 1630 |
| Db | 521 | AlaGlyValThrSerTrpGlyThrGlyCyseIyIyIlnArgIeuIySProGlyValTyraThr | 540 |
| QY | 1631 | AAAGTGCACGAAGTCTTCTCCCTGGATTTTACACGAAGATGAGAGACGAGTGCCTTACGA | 1690 |
| Db | 541 | LysValThrGluValIeuProItrpIleTyraSerIySwetGluSerGluValaArgPheArg | 560 |
| QY | 1691 | AAATCC 1696 | |
| Db | 561 | LysSer 562 | |
| RESULT 3 | | | |
| US-10-806-370-12 | | | |
| Sequence 12, Application US/10806370 | | | |
| Publication No. US20040209327A1 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: Xiao, Yonghong | | | |
| APPLICANT: Gedrich, Richard | | | |
| TITLE OF INVENTION: Regulation of Human Transmembrane Serine | | | |
| FILE REFERENCE: 02973_00035 | | | |
| CURRENT APPLICATION NUMBER: US/10/806,370 | | | |
| CURRENT FILING DATE: 2004-03-23 | | | |
| PRIOR APPLICATION NUMBER: US/09/879,792 | | | |

RESULT 3

US-10-806-370-12
Sequence 12, Application US/10806370
Publication No. US20040209327A1
GENERAL INFORMATION:
Applicant: Xiao, Yonghong
Applicant: Gedrich, Richard
Title of Invention: Regulation of Human Transmembrane Serine
Protease
File Reference: 02973_00035
Current Application Number: US/10/806,370
Current Filing Date: 2004-03-23
Prior Application Number: US/09/879,792

Publication No. US20040001801A1
GENERAL INFORMATION:
APPLICANT: Edwin L. Madison
APPLICANT: Joseph Edward Sempke
APPLICANT: George P. Vlasuk
APPLICANT: Scott Jeffrey Kemp
APPLICANT: Mallareddy Komandla
APPLICANT: Daniel Vanna Siev
TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
FILE OF INVENTION: Thereof
FILE REFERENCE: 24745-1611
CURRENT APPLICATION NUMBER: US/10/156,214A
CURRENT FILING DATE: 2002-05-23
NUMBER OF SEQ. ID NOS: 611
SOFTWARE: PASCSEQ for Windows Version 4.0
SEQ ID NO 24
LENGTH: 562
TYPE: PRN
ORGANISM: homo sapien
FEATURE:
NAME/KEY: protease domain of endothelialase 2
LOCATION: (321)..(562)
US-10-156-214A-24

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 5,14e-174 | Length: | 562 |
| Score: | 2991.00 | Matches: | 561 |
| Percent Similarity: | 99.8% | Conservative: | 0 |
| Best Local Similarity: | 99.8% | Mismatches: | 1 |
| Query Match: | 89.7% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-10-806-370-11 (1-1748) x US-10-156-214A-24 (1-562)

11 ATGAGAGGAGCAGACCGGAGATGATCTTCAGACCAAGAACCTTCAGCTGAGCATCT 70
1 MeegluargrapsrshisgllyAsnAlaserProhlarhgrhProsserAlaglyAlaser 20
71 CCAAGCCAGGATTCAGAGCTGGAGACCTTCAGAGCCGGGATTCAGACCCAGGATCT 130
21 ProhlarhAlaserProhlarhglYThrProhglYArghAlaserProhlarhAlaser 40
131 CCAAGCCAGGATTCAGAGCTGGAGACCTTCAGAGCCGGGATTCAGACCCAGGATCT 190
41 ProhlarhAlaserProhlarhglYThrProhglYArghAlaserProhlarhAlaser 60
191 CCAAGCTGATACCTTCAGAGCCGGGATTCAGAGCCGGGATTCAGACCCAGGATCT 250
61 ProhlarhglYThrProhglYArghAlaserProhlarhglYThrProhlarhAlaser 80
251 CCAAGCCGGGATTCAGAGCTGGAGACCTTCAGAGCTTCATTCAGAGCTTCAGAGCTCA 310
81 ProhlarhAlaserProhlarhAlaserleuAlaserleuSerleuSerleuSerleuSerleu 100
311 TCCGCGAGGATTCAGAGCTGGAGACCTTCAGAGCTTCATTCAGAGCTTCATTCAGAGCA 370
101 SerhlarhSerhAlaserValThrSerProhlarhglYThrValAlaArgAlaThr 120
371 CCAAGTGGAGCTGATACCTTCAGATTCATTCCTTCAGAGCTTCAGAGCAAGGAGCC 430
121 ProhlarhglYThrAlaProhlarhglYThrSerleuSerleuSerleuSerleuSerleu 140
431 ACCAGGAGAGCCAGGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCA 490
141 ThrArghglYThrProhglYThrSerleuProhlarhglYThrArghglYThrValAla 160
491 CTACCGCTTCAGAGCTGGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAG 550
161 LeuProhlarhglYThrValAlaSerleuSerleuSerleuSerleuSerleuSerleu 180
551 TTCAGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAG 610
181 PhehlarhglYThrValAlaSerleuSerleuSerleuSerleuSerleuSerleuSerleu 200

611 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
201 LyhAlaAlaValArgCyAspArglyValAlaAspCyAspLeuSerleuSerleuSerleuSerleu 220
671 TGGGTGAGCTTTCAGAGCTGAGCAAGAGCTTCCTTCAGAGCTTCAGAGCTTCAGAGCTTCAG 730
221 CyValArghPhehlarhglYThrArghglYThrSerleuSerleuSerleuSerleuSerleu 240
731 TGGCTTCAGATTCAGAGCTGAGCAAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 790
241 ThrleuProhlarhglYThrValAlaSerleuSerleuSerleuSerleuSerleuSerleu 260
791 CTGAGCTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
261 LeuGlYThrPhehlarhglYThrValAlaSerleuSerleuSerleuSerleuSerleuSerleu 280
851 TTCATTCAGATTCAGAGCTGAGCAAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 910
281 PheSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleu 300
911 TCCAGCGATATATCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 970
301 SerGlYThrPhehlarhglYThrValAlaSerleuSerleuSerleuSerleuSerleuSerleu 320
971 ATGAGTGGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030
321 LLevalAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 340
1031 GGCAGC 1090
341 GlYThr 360
1091 CACTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
361 HieCyPhePheValThrArghglYThrValAlaSerleuSerleuSerleuSerleuSerleu 380
1151 AGCAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 1210
381 SerhlarhglYThrProhglYThrValAlaSerleuSerleuSerleuSerleuSerleuSerleu 400
1211 TACAGGATGAGAGGAGCAGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 1270
401 TyrThrAspGlYThrAspArglyThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 420
1271 CTGCT 1330
421 LeuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleu 440
1331 GAGACCTGCTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 1390
441 GluThrCyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 460
1391 TTCCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1450
461 PheleuArglyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 480
1451 TATGACATTCATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAG 1510
481 TyrAspSerleuThrProhlarhglYThrValAlaSerleuSerleuSerleuSerleuSerleu 500
1511 TCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1570
501 SerCySerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleuSerleu 520
1571 GCAGTGTTCAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1630
521 AlaglyValThrSerleuThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 540
1631 AAAGTGCAGAGCTTCCTGCTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCT 1690
541 LyValThrThrValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 560

QY 1691 AATCC 1696
 Db 561 LysSer 562

RESULT 5

US-10-916-758-94
 ; Sequence 94, Application US/10916758
 ; Publication No. US20050180977A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nixon, Andrew
 ; APPLICANT: Madison, Edwin L.
 ; TITLE OF INVENTION: ENDOTHELINASE-2 LIGANDS
 ; FILE REFERENCE: 10280-065001
 ; CURRENT APPLICATION NUMBER: US/10/916,758
 ; PRIORITY FILING DATE: 2004-08-12
 ; PRIORITY FILING DATE: 2003-11-14
 ; PRIORITY FILING DATE: 2003-08-14
 ; NUMBER OF SEQ ID NOS: 113
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 94
 ; LENGTH: 562
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-10-916-758-94

Alignment Scores:

Pred. No.: 5,14e-174 Length: 562
 Score: 2991.00 Matches: 561
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 89.7% Indels: 0
 Gaps: 0

US-10-806-370-11 (1-1748) x US-10-916-758-94 (1-562)

QY 11 ATGAGAGAGGACAGCCAGGGAATGCATCTCCAGCAAGAACCTTCAGCTGAGCATCT 70
 Db 1 MetGlnArgSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAGCCAGGATCTCCAGCTGGAGACACTCCAGCCGGGATCTCCAGCCAGGATCT 130
 Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAGCCAGGATCTCCAGCTGGAGACACTCCAGCCGGGATCTCCAGCCAGGATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAGCTGGTACATCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGCCAGGATCT 250
 Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAGCCCGGGATCTCCGGGCTTGGCATCACTTTCAGAGTCTCTCATCCGGAGGTCATCA 310
 Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
 QY 311 TCCGCCAGGTGAGCTCGGTGACAACTCCCAAGTTCACTTGGGGAGGGGCAAGAAG 370
 Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
 QY 371 CCAGTGGGGGTGTACCCATCCATCATCTCTGCCAGTGCAGACCAAGAGGAGCC 430
 Db 121 ProValGlyAlaValProLeuArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGACCCAGGTACAGAGCTGCCCAAGTTCACTTGGGGAGGGGCAAGAAG 490
 Db 141 ThrArgGlnSerProGlyThrSerLeuProLysPheThrTyrParGlnGlnLysGln 160
 QY 491 CTACCGCATCGGGTGGCTCTCTCATTTGGCCCTGAGTTTGAGTCATCATCTCTC 550
 Db 161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIleLeu 180

QY 551 TTCAGATTCTGGACAGGCGCACACAGGATCAGGTACAGAGACAGAGGAGAGCTGTCC 610
 Db 181 PheGlnPheThrPrlGlnGlyHisThrGlyIleArgTyrLysGlnArgGlnSerCysPro 200
 QY 611 AAGCAAGCTGTTCGCTGTGACCGGGGTGTGGACTTCGCAAGCTGAAGACTGACAGCTGGC 670
 Db 201 LysHisAlaValArgCysArgGlyValValAspCysLeuValSerAspGlnLeuGly 220
 QY 671 TGCCTGAGATTGACTGGGACAAAGTCTCTGTTAAATCTACTCTGGGCTCTCCATCAG 730
 Db 221 CysValArgPheAspThrPAspLysSerLeuLeuValIleTyrSerGlySerSerHisGln 240
 QY 731 TGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAGAAGACTGCCAGAG 790
 Db 241 TrpLeuProIleCysSerSerAspThrPAspAspPertyrSerGlnLysThrCysGlnGln 260
 QY 791 CTGGGTTTCAGACAGTGTCAACCGGACCAAGAGGTTGCCCAAGGATTTTGCCAAACAG 850
 Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGluValAlaHisArgAspPheAlaSerSer 280
 QY 851 TTCTCAATCTTGAGATACAACTCCACCATCCAGAGAAAGCTCCACAGGTCTGAATGCCCT 910
 Db 281 PheSerIleLeuArgTyrAspSerThrIleGlnGlnSerLeuHisArgSerGlnCysPro 300
 QY 911 TCCAGCGGTATATCTCCCTCCAGTGTTCCTCACTGCCACTGAGGCGCATGACCGGCGG 970
 Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
 QY 971 ATCGTGGAGGGGCGCTGGGCTGGATAGCAAGTGGCTTGGCAAGTGAAGCTTCACATTC 1030
 Db 321 IleValIleGlyAlaLeuAlaSerAspSerTyrThrProIleGlnValSerLeuHisPhe 340
 QY 1031 GGCACACCCACATCTGTGAGAGGACCGCTCATGACGCCACAGTGGTGTCTACTCGCCG 1090
 Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnThrValIleThrAlaAla 360
 QY 1091 CACTGCTTCTTCTGACCCCGGAGAGAGTCTTGGAGGGCTGAAAGGTGTACCGGGGACC 1150
 Db 361 HisCysPhePheValThrArgGlnLysValLeuGlnIleGlyThrValIleGlyThr 380
 QY 1151 AGCAACCTGACCCAGTTCGCTGGAGGAGCTCCATTCGCCGATCATCATCAAGCAAT 1210
 Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleAsnSerAsn 400
 QY 1211 TACACCATGAGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGAC 1270
 Db 401 TyrThrAspGlnGluAspAspTyrAspIleAlaLeuLeuAspArgLeuSerLysProLeuThr 420
 QY 1271 CTGTCCGCTCAGATCCACCTGCTGCTGCCCTCCCATGATGAGACAGACTTTAGCTCAT 1330
 Db 421 LeuSerAlaHisIleHisAspProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
 QY 1331 GAGACCTGCTGGATCAGAGGCTTGGAGAGACAGGAGAGACAGATGACAAGATCCCCC 1390
 Db 441 GlnThrCysThrIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
 QY 1391 TTCCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTGTGTC 1450
 Db 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
 QY 1451 TATGACAGTTTACCTTACCCCAAGATGATGTGTCTGGGACCTTGTGGGGGCGAGAGAC 1510
 Db 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
 QY 1511 TCCGTGCAAGGAGACAGCGGGGGGCTCTGTCTGTGAGAGAGAACACCGGTGTACTGCTG 1570
 Db 501 SerCysGlnGlnLysAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTyrPtyLeu 520
 QY 1571 GCAAGTGTCAACAGCTGGGCGACAGGCTGTGCGCCAGAGAAACAACTGTGTGTACACC 1630
 Db 521 AlaGlyValThrSerThrPrlGlyThrGlyCysGlyGlnArgAsnLysProGlyValIleThr 540
 QY 1631 AAAGTGAACAGAGTTCTTCCCTGGATTTCACAGCAAGATGGAAGCGAGGTGCGATTGACA 1690

Db 541 LysalThrGluValLeuProTPrIleTyrSerIleGluSerGluValArgPheIle 560
QY 1691 AAATCC 1696
Db 561 LysSer 562

RESULT 6
US-10-156-214A-26
; Sequence 26, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Semple
; APPLICANT: George P. Vlasuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Siev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Us
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156, 214A
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: protease domain
; LOCATION: (321)..(688)
US-10-156-214A-26

Alignment Scores:
Pred. No.: 2,69e-172 Length: 688
Score: 2963.00 Matches: 554
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 88.9% Indels: 0
Gaps: 0
DB: 4

US-10-806-370-11 (1-1748) x US-10-156-214A-26 (1-688)

QY 11 ATGAGAGGAGACACCGGAGATGATCTTCACCAAGACACTTCACTGAGCATCT 70
Db 1 MetGluArgSerSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGATTCAGCTGGAGACACTTCGAGCGGAGCATCTCCAGCCAGGACATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCAGGATTCAGCTGGAGACACTTCGAGCGGAGCATCTCCAGCCAGGACATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGGTACCTCCAGGCGGAGCATCTTCAGGCGGAGCATCTCCAGCCAGGACATCT 250
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAGCCCGGAGCATCTCGGCTGGAGCATCTTCAGGCTGCTCATCCGAGGAGTCA 310
Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerSerGlyArgSerSer 100
QY 311 TCCGCGAGGTCAGCTCGGTGACAACTCCCAACCAAGAGTGAACCTTGTAGAGACA 370
Db 101 SerIleArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCAAGTGGGGCTGTAACCATCCGATCATCTCTGCGAGGTGACAGCAACCAAGAGGCC 430
Db 121 ProValGlyAlaValProIleArgSerSerProIleArgSerAlaProIleThrArgAla 140
QY 431 ACCAGGAGAGCCCAAGTACAGAGCTGCCAAGTTACCTGCGGAGAGGCCAGAGAG 490

Db 141 ThrArgIleSerProGlyThrSerIleuProIleuPheThrIleArgGluGlnIleuGln 160
QY 491 CTACCGCTCATCGAGTGCCTGCTCTCTCATTCGCTGCTGTTGCTCATATCTCTC 550
Db 161 LeuProIleuIleGlyCysValLeuIleuIleAlaLeuValIleSerIleIleIleu 180
QY 551 TTCAGTTCGGAGGGCCACAGAGGATCAGGATGACAGAGAGAGAGAGAGAGAGAG 610
Db 181 PheGlnPheThrGlnGlyIleThrGlyIleArgGlyIleuGlnIleuGlnIleuGln 200
QY 611 AACGACGCTTGGCTGTGACAGGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 670
Db 201 LysHisAlaValArgCysArgGlyValAlaAspCysIleuIleuSerArgGluGly 220
QY 671 TGGGTGAGTTTGAATCGGACAAAGTCTGCTTAAATTAATCTGCTGGCTCTCCATAG 730
Db 221 CysValArgPheArgPheArgPheArgPheArgPheArgPheArgPheArgPheArg 240
QY 731 TGGCTTCCCATCTGTAGACAACTGGAATGACTCTTACAGAGAGAGAGAGAGAGAG 790
Db 241 TyrLeuProIleCysSerSerIleThrAspAspSerTyrSerGluIleuIleuGln 260
QY 791 CTGGGTTTCGAGAGTGTCAACCGACAAACGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
Db 261 LeuGlyPheGluSerAlaHisArgThrThrGluValAlaHisArgPheAlaHisSer 280
QY 851 TTCTCAATCTTGAGATCAATCCACCATCCAGAAAGCTCCACAGTGTGAATGCTCT 910
Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnIleuSerIleuIleuIleuIleuIleu 300
QY 911 TCCGAGGATATATCTCCCTCCAGTGTCTCCAGTGGAGAGAGAGAGAGAGAGAGAG 970
Db 301 SerGlnArgTyrIleSerIleuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
QY 971 ATCGTGGAGAGGCGGCTGCTGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1030
Db 321 IleValGlyIleAlaAlaAlaSerAlaSerIleThrProIleGlnIleuIleuIleuIleu 340
QY 1031 GGACACACCAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090
Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnThrValIleuThrAlaAla 360
QY 1091 CACTGCTTCTTCTGAGACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150
Db 361 HisCysPhePheValThrArgGluIleuValLeuGlnIleuIleuIleuIleuIleu 380
QY 1151 AGCACTGACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1210
Db 381 SerAsnLeuHisGlnLeuProGlnAlaIleSerIleAlaGlnIleuIleuIleuIleuIleu 400
QY 1211 TACACGATGAG 1270
Db 401 TyrThrAspGluGluAspArgTyrAspIleAlaIleuMetArgLeuSerIleuProLeuThr 420
QY 1271 CTGTCGCTCAATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1330
Db 421 LeuSerIleHisIleHisProAlaCysLeuProMetHisGlyGlnIleuIleuIleuIleuIleu 440
QY 1331 GAGACGCTGAGATCAAG 1390
Db 441 GluThrCysTrpIleThrGlyPheGlyIleuThrArgGluThrAspArgIleuThrSerPro 460
QY 1391 TTCCTCCGAG 1450
Db 461 PheLeuArgGluValGlnValAlaSerIleAspPheIleuIleuIleuIleuIleuIleu 480
QY 1451 TATGACAGTTACCTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1510
Db 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaIleuAspLeuArgGlyGlyIleuAsp 500
QY 1511 TCTGTCAGGAG 1570

Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnArgTrpTyrLeu 520
Qy 1571 GCAGGTGTACACGCTGGGGACAGAGCTGTGGCCAGAAACAAACCTGGTGTGACACC 1630
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
Qy 1631 AAAGTGAAGAAAGTTCTTCCCTGGATTTTACGACAAGATGAGACCGAGTCCGA 1684
Db 541 LysValThrGluValLeuProTrpIleTyrSerLysMetGluAsnArgAlaGln 558

RESULT 7

US-10-916-758-2
; Sequence 2, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELINASE-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-916-758-2

Alignment Scores:

Pred. No.: 2,696-172 Length: 688
Score: 2963.00 Matches: 554
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 0
Query Match: 88.9% Indels: 0
Gaps: 5

US-10-806-370-11 (1-1748) x US-10-916-758-2 (1-688)

Qy 11 ATGAGAGAGGACAGACCGGGAATGATCTTCAGACAAGAACCTTGAGTGAACATCT 70
Db 1 MetGluArgAspSerHisGlyAsnAlaSerProIleArgThrProSerAlaGlyAlaSer 20
Qy 71 CCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGCCAGGATCT 130
Db 21 ProIleGlnAlaSerProIleGlyThrProProGlyArgAlaSerProIleGlnAlaSer 40
Qy 131 CCAGCCAGGATCTCCAGCTGGGACACCTCCGGGCGGGATCTCCAGCCAGGATCT 190
Db 41 ProIleGlnAlaSerProIleGlyThrProProGlyArgAlaSerProIleGlnAlaSer 60
Qy 191 CCAAGTGTATACCTCCAGGCGGGGATCTTCAGGCGGGATCTCCAGCCAGGATCT 250
Db 61 ProIleGlyThrProProGlyArgAlaSerProGlyArgAlaSerProIleGlnAlaSer 80
Qy 251 CCAGCCGGGATCTCCGGCTCTGGCATCACTTTCCAGGTCTTCAATCCGGAGGTCACT 310
Db 81 ProIleArgAlaSerProIleAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
Qy 311 TCCGCCAGGTGAGCTGGTGAACAACCTCCCAACAGAGGTGACTTGTAGAGACA 370
Db 101 SerIleArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
Qy 371 CCAAGTGGGGGTGTATCCATCCATCATCTCTGCAAGTGCAGACCAAGCAAGGGCC 430
Db 121 ProValGlyAlaValProIleArgSerSerProIleArgSerAlaProIleThrArgAla 140
Qy 431 ACCAGGAGAGACCCAGGTACGAGCTGCGCAAGTTCACTTGCGGAGGGCCAGAGACG 490

Db 141 ThrArgLysSerProGlyThrSerLeuProLysPheThrTrpArgGluGlnLysGln 160
Qy 491 CTACCGCTCATCGGGGTGCGTCTCTCTCATTCAGCCCTGGGTGGTTGCTCATCATCTC 550
Db 161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValAlaSerLeuIleIleLeu 180
Qy 551 TTCAGTTCTGGACAGGGCCACACAGGGATCAGGTACAAAGACAGAGAGAGCTGTCC 610
Db 181 PheGlnPheTrpGlnGlyIleThrGlyIleArgTyrLysGlnGlnArgLysSerCysPro 200
Qy 611 AAGACGCTGTTCGCTTCGACGAGGCTGGTGGTGAAGTTCGACGAGTTCAGCTGGCC 670
Db 201 LysHisAlaValAlaGlyCysAspGlyValValAspCysLysLeuLysSerAspGlyLeuGly 220
Qy 671 TGCGTGAGGTTTGACTGGGACAAGTCTCTGCTTAAATCTACTTGAGTCTGCTCCATCAG 730
Db 221 CysValArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHisGln 240
Qy 731 TGCCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTAGAGAAACCTGCCACAG 790
Db 241 TrpLeuProIleCysSerSerAsnTrpAsnAspSerTyrSerGlyLysThrCysGlnGln 260
Qy 791 CTGGGTTTCGAGAGTGTCAACCGGACAAACGAGTTCGCCAGAGGATTTGCCAAGAC 850
Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGluValAlaHisArgAspPheAlaAsnSer 280
Qy 851 TTCTCATCTTGAGATACAACTCCACCATCCAGGAAGACCTCCACAGGTCTGAATGCCCT 910
Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuHisArgSerGlyCysPro 300
Qy 911 TCCAGCGGTATATCTCCCTCCAGTGTTCCTCACTGCGGACTGAGGCGCATGACCGGCGG 970
Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
Qy 971 ATCTGGAGAGGGCGCTCGGCTCGGATAGCAAGTGGCTTGCAAGAGTGTCCATCTC 1030
Db 321 IleValIleGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHisPhe 340
Qy 1031 GGCACACCAACATCTGTGAGAGGACCGCTCATTTGACCCCGAGTGGGTCTCACTGCCGC 1090
Db 341 GlyThrThrHisIleIleCysGlyGlyThrLeuIleAspAlaGlnTrpValIleuThrAlaAla 360
Qy 1091 CACTGCTTCTTCTGACCCCGGAGAGAGTCTTGAGGGCTGGAAGGTGTACCGGGACCC 1150
Db 361 HisCysPhePheValThrArgGlyLysValLeuGlnIleTyrLysValTyrAlaGlyThr 380
Qy 1151 AGCAACCTGACACCAAGTGTGAGGAGCGCTCCATTCGAGATCATCATCAACAGCAAT 1210
Db 381 SerAsnLeuHisGlnLeuProGluAlaAlaSerIleAlaGluIleIleIleAsnSerAsn 400
Qy 1211 TACACGATGAGAGAGACGACTATGACATGCGCCTCATGCGGCTGTCCAAAGCCCTGACC 1270
Db 401 TyrThrAspGluGluAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
Qy 1271 CTGTGCTCATCATCACCTGCTTGCCTCCCATGATGACAGACCTTTAGCTCAAT 1330
Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
Qy 1331 GAGACCTGCGAGTACAGAGCTTTGGCAAGACCAAGGAGACAGATGACAGATCCCCC 1390
Db 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysThrSerPro 460
Qy 1391 TTCCTCCGAGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGACTACTTGATC 1450
Db 461 PheLeuArgGluValGlnValAlaLeuLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
Qy 1451 TATGACAGTTAATCTTACCCCAAGATATGTGTGCTGGGACCTTCTGGGGCGAGAGAC 1510
Db 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
Qy 1511 TCCGCGCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGAGAAACAACCGCTGTACTCTG 1570
Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnArgTrpTyrLeu 520

QY 1571 GCAGGTGTCACGAGTGGGGGACAGCGCTGTGGCCAGAGAAACCTGTGTGTACACC 1630
DB 521 AAGTGTGTCACGAGTGGGGGACAGCGCTGTGGCCAGAGAAACCTGTGTGTACACC 1630
QY 1631 AAGTGTGTCACGAGTGGGGGACAGCGCTGTGGCCAGAGAAACCTGTGTGTACACC 1630
DB 541 AAGTGTGTCACGAGTGGGGGACAGCGCTGTGGCCAGAGAAACCTGTGTGTACACC 1630
RESULT 8
US-10-353-690-100
Sequence 100, Application US/10353690
Publication No. US20030215840A1
GENERAL INFORMATION:
APPLICANT: Logan, Thomas Joseph
APPLICANT: Chun, Myoung
APPLICANT: Galvin, Katherine M.
APPLICANT: Healy, Aileen
APPLICANT: Acton, Susan L.
APPLICANT: Donoghue, Mary
APPLICANT: Stagliano, Nancy
APPLICANT: Perodini, Jacqueline
APPLICANT: Rodriguez-Way, Amelie
TITLE OF INVENTION: Methods and compositions for treating
TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
TITLE OF INVENTION: 283, 2554, 9464, 17799, 26666, 43848, 32135, 12208, 2914,
TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
FILE REFERENCE: MP102-0181RMONIM
CURRENT FILING DATE: 2003-01-29
PRIOR APPLICATION NUMBER: US/10/353,690
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 60/353,224
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 60/364,529
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: 60/373,861
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/376,287
PRIOR FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 60/388,080
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: 60/390,971
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/394,130
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: 60/394,797
PRIOR FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/404,904
PRIOR FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/405,450
PRIOR FILING DATE: 2002-08-23
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 126
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 100
LENGTH: 581
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-353-690-100
Alignment Scores:
Pred. No.: 7,03e-172 Length: 581
Score: 2956.00 Matches: 553
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 88.7% Indels: 0
DB: 4 Gaps: 0
US-10-806-370-11 (1-1748) x US-10-353-690-100 (1-581)

QY 11 ATGGAGAGGACAGACCGGGAATGATCTCCAGACAGAAACCTTCAGTGGAGCATCT 70
DB 1 MetcUnrghspsrhstgslgysanlaserProlaarhtrProserlaaglylaser 20
QY 71 CCAGCCGAGCATCTCCAGTGGAGACCTTCAGACCGGGGATCTCCAGCCGAGCATCT 130
DB 21 ProlaaglnlaserProlaaglylThrProProgllylrglaserProlaaglnlaser 40
QY 131 CCAGCCGAGCATCTCCAGTGGAGACCTTCAGACCGGGGATCTCCAGCCGAGCATCT 190
DB 41 ProlaaglnlaserProlaaglylThrProProgllylrglaserProlaaglnlaser 60
QY 191 CCAGCCGAGCATCTCCAGTGGAGACCTTCAGACCGGGGATCTCCAGCCGAGCATCT 250
DB 61 ProlaaglylThrProProgllylrglaserProlaaglnlaserProlaaglnlaser 80
QY 251 CCAGCCGAGCATCTCCAGTGGAGACCTTCAGACCGGGGATCTCCAGCCGAGCATCT 310
DB 81 ProlaarghlaserProlaaglnlaserProlaaglnlaserProlaaglnlaser 100
QY 311 TCCGCGAGTTCAGCTTCGAGCAACCTCCCAACAGAGTTCCTTTAGAGCAACA 370
DB 101 SerlaarghlaserProlaaglnlaserProlaaglnlaserProlaaglnlaser 120
QY 371 CCAGTGGGGGCTGATCCATCCGATCCATCTTCGAGTGGAGACAGACAGAGGCGC 430
DB 121 ProvalglalvalProlaaglnlaserProlaaglnlaserProlaaglnlaser 140
QY 431 ACCAGGAGAGCCGAGTTCAGACCTTCAGACCTTCAGACCTTCAGACCTTCAGAC 490
DB 141 ThrarglulserProgllylThrSerleuProleuPheleuThrParglulglulser 160
QY 491 CTACGCTGATCGGAGTGTGCTCTCTCTCAATGCTGCTGCTGCTGCTGCTGCTGCT 550
DB 161 LeuProleuilegllylcyvalleuileuileuileuileuileuileuileuileu 180
QY 551 TTCCAGTTCGAGCGGCGCACAGGAGTTCAGAGTTCAGAGGAGGAGGAGGAGTTC 610
DB 181 PheglulserProgllylThrSerleuProleuPheleuThrParglulglulser 200
QY 611 AAGCAGCTGCTGCTGAGACCGGGGATGAGTTCAGACCTTCAGACCTTCAGAC 670
DB 201 Lyslshsalavalagcyshapgllylvalalpccylsyleuylserleuileuileu 220
QY 671 TGCGTGAGTTCAGTGGAGCAACCTCTGCTTAAATCTGCTGCTGCTGCTGCTGCT 730
DB 221 CysvalarghlaserProlaaglnlaserProlaaglnlaserProlaaglnlaser 240
QY 731 TGGCTTCCATCTGAGCAGCACTGGAATGATCTCTAGAGAAACCTTCAGACGAG 790
DB 241 TTPleuProleuilegllylThrSerleuProleuPheleuThrParglulglulser 260
QY 791 CTGAGTTCGAGTTCAGTGGAGCAACCTTCAGACCTTCAGACCTTCAGACCTTCAG 850
DB 261 Leuileuileuileuileuileuileuileuileuileuileuileuileuileu 280
QY 851 TTCTCAATCTTGAATCAACTCCAGCATTCAGAAAGCTTCAGAGTTCGAAATGCT 910
DB 281 Pheleuileuileuileuileuileuileuileuileuileuileuileuileu 300
QY 911 TCCAGCGGATATCTCCCTTCAGTTCCTTCAGTTCCTTCAGTTCCTTCAGTTCCT 970
DB 301 Serlinalrglylserleuileuileuileuileuileuileuileuileuileu 320
QY 971 ATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1030
DB 321 IlevalgllylalaalaserProlaaglnlaserProlaaglnlaserProlaagln 340
QY 1031 GGACACCAACATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1090
DB 341 GlyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 360

| | | | |
|----|------|---|------|
| QY | 1091 | CACGCGTCTCTTCGGAGACCCCGGAGAAAGTCTCGAGGGCTGGAGAGGTGACCGGGGAC | 1150 |
| Db | 361 | HiccyapPepPevAlThrArgIuValLeuGIuGlyTrpIyValTyAlaGIlyThr | 380 |
| QY | 1151 | AGCAACCTGCACCAAGTGGCTGAGGAGAGCTCCATGTGGCGAGATCATCAACAGCAAT | 1210 |
| Db | 381 | SerLeuLeuH1aGIuLeuProGIuAlaSerIleAlaGIuIleIleIleAsnSerAsn | 400 |
| QY | 1211 | TACACCGATGAGAGAGAGCACTATGACATGCGCCCTCATCGGGCTGTCCAAAGCCCTGAC | 1270 |
| Db | 401 | TyrThrAspGIuGIuAspAspTyAspIleAlaLeuMetArgLeuSerIyProLeuThr | 420 |
| QY | 1271 | CTGTCGGCTGCATCCACCCCTGCTTGGCTCCCGCATGCATGGACAGACCTTTAAGCTCAAT | 1330 |
| Db | 421 | LeuSerAlaHisIleHisIleProAlaCysLeuProMetHisGIuGlnThrPheSerLeuAsn | 440 |
| QY | 1331 | GAGACCTGCTGGATCACAGGCTTTGGCAAGCAAGGAGACAGATGACAAAGATCCGCC | 1390 |
| Db | 441 | GIuThrCysTrpIleThrGIuPheGIuIyysThrArgIuThrAspAspIysthSerPro | 460 |
| QY | 1391 | TTCTCCGGAGGAGTGACAGTCAATCTCATCGACTTCAAGAAATGCATGACTTACTTGGTC | 1450 |
| Db | 461 | PheLeuArgIuValGIuValIleAsnLeuIleAspPheIyIyScyAsnAspTyLeuVal | 480 |
| QY | 1451 | TATATACAGTTACCTTACCCCAAGATATATGTGTCTGGGACCTTCTGTGGGGGACAGAC | 1510 |
| Db | 481 | TyrAspSerTyLeuLeuThrProAlaMetMetCysAlaGIyAspLeuArgGIyIyArgAsp | 500 |
| QY | 1511 | TTCGACCAAGGAGACAGCGGGGGGCTTCTGTGTGGAGCAACAACCGCTGGTACCTG | 1570 |
| Db | 501 | SerCysGIuGIuAspSerGIyGIyProLeuValCysGIuGIuAsnAsnAspTrpTyLeu | 520 |
| QY | 1571 | GCAGGTGCACAGACTGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACAC | 1630 |
| Db | 521 | AlaGIyAlaThrSerTrpGIyThnGIyCysGIyGlnArgAsnIySerGIyValIyThr | 540 |
| QY | 1631 | AAAGTGACAGAACTTCTCCTGGATTTACACAGATGAG | 1672 |
| Db | 541 | IyValAlaThrGIuValLeuProTrpIleTySerIyMetGIu | 554 |

| Alignment Scores: | | | |
|---|---------|---|-----|
| Pred. No.: | 1e-171 | Length: | 593 |
| Score: | 2953.50 | Matches: | 555 |
| Percent Similarity: | 99.1% | Conservative: | 0 |
| Best Local Similarity: | 99.1% | Mismatches: | 0 |
| Query Match: | 88.6% | Indels: | 5 |
| DB: | 4 | Gaps: | 1 |
| US-10-806-370-11 (1-1748) x US-10-428-275-120 (1-593) | | | |
| QY | 8 | ACCATGGAGGAGCAGCACCGGGAATGCATCTCCAGACAGAAACACCTTCAGTGGAGCA | 67 |
| DB | 4 | ThrmecGluhrghspsberthiegllyenmlaserProAlaArgThrProSerAlaGlyAla | 23 |
| QY | 68 | TTCTCCAGCCAGGCATCTCCAGTGGGACACTCCAGGCCGGGCATCTCCAGCCAGCA | 127 |
| DB | 24 | SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla | 43 |
| QY | 128 | TTCTCCAGCCAGGCATCTCCAGTGGGACACTCCGGGGCGGGGATCTCCAGCCAGCA | 187 |
| DB | 44 | SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla | 63 |
| QY | 188 | TTCTCCAGTGGTACACCTCCAGGCGGGCATCTCCAGGCCG-----GCA | 232 |
| DB | 64 | SerProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAla | 83 |
| QY | 233 | TTCTCCAGCCAGGCATCTCCAGCCGGGCATCTCGGCTTCGGCATCTTCAGAGTCC | 292 |
| DB | 84 | SerProAlaGlnAlaSerProAlaArgAlaSerProAlaGlnAlaSerLeuSerArgSer | 103 |
| QY | 293 | TCAATCGGGAGGTCATCATCCGCGCAGGTACGCTCGTGACAACTCCGCCAACAGAGTG | 352 |
| DB | 104 | SerSerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgVal | 123 |
| QY | 353 | TACCTTGTTAGACCAACACAGTGGGGGTGTACCCATCCGATCATCTCTCCAGAGTCA | 412 |
| DB | 124 | TyrLeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSer | 143 |
| QY | 413 | GCACACAGCAACCGGGCCACCGGAGAGCCAGGTACGAGCTCCGCCAGTTCACTGG | 472 |
| DB | 144 | AlaProAlaThrArgAlaThrArgIleuSerProGlyThrSerLeuProLysPheThrTyr | 163 |
| QY | 473 | CGGAGGGCCAGAGAGAGTACCGGTCATCGGGTGCAGTCTCTCATTTGAGCCTGGTG | 532 |
| DB | 164 | ArgGlnuGlyGlnuSerGlnuProLeuIleGlyCysValLeuLeuLeuIleAlaIleVal | 183 |
| QY | 533 | GTTCGGTCATCATCTCTTCAGTTCTGGACGGGCCACACAGGGATCAGGTACAGAG | 592 |
| DB | 184 | ValSerLeuIleIleLeuPheGlnPheThrGlnGlyIleThrGlyIleArgTyrGlyGlu | 203 |
| QY | 593 | CAGAGGAGAGAGTGTCCCAAGACCGCTGTCCCTGTGACGGGGTGGGAGATTGGAACGTG | 652 |
| DB | 204 | GluArgGlnuSerCysArgProLysIleAlaValArgCysAspGlyValValAspCysValLeu | 223 |
| QY | 653 | AAGAGTACAGAGCTGGGCTGCTGAGAGTTTGACTGGACAAGTCTGTGTTAAATATAC | 712 |
| DB | 224 | LysSerAspGlnuGlnuGlyCysValArgPheAspTyrPheLysSerLeuLeuLysIleTyr | 243 |
| QY | 713 | TTTGGGCTCCCATCAGTGGCTTCCCATCTGTACAGACCACTGGATGACTCTTACTCA | 772 |
| DB | 244 | SerGlySerSerHisGlnThrLeuProIleCysSerSerAsnThrPheAspSerTyrSer | 263 |
| QY | 773 | GAGAAAGACTGCGACAGAGCTGGGTTTCCAGAGTGTCAACCGGACAACTGAGATTGCCAC | 832 |
| DB | 264 | GlnuLysThrCysGlnuGlnuGlyPheGlnuSerIleAlaIleAspGlyThrGlnuValAlaHis | 283 |
| QY | 833 | AGGAGTTTGCACAGCTTTCATCTTGAGATCACTCAACTCCAGCATCAGAGAAGCTC | 892 |
| DB | 284 | ArgPhePheAlaAsnSerPheSerIleLeuArgTyrAsnSerThrIleGlnuGlnuSerLeu | 303 |
| QY | 893 | CACAGGTTGGAATGCCCTTCCAGCGGATATATCTCCCTCAGTGTTCCCATCGGAGCTG | 952 |

Db 304 HieArSerGIuCyProSerGIuArgTyrIleSerLeuGIuIncysSerHisCysGlyLeu 323
QY 953 AGGGCCATGACCGGGCGGATCGTGGGAGGGCGCTGGCCCTCGATTACAGTGGCCCTGG 1012
Db 324 ArgAlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTrpProTrp 343
QY 1013 CAAGTGAAGTGTGCACTTCGGACACCCACATCTGTGGAGGACGCTCATTTGACGCCAG 1072
Db 344 GluValSerLeuHisPheGlyThrThrHisIleLeuGlyGlyThrLeuIleAspAlaGln 363
QY 1073 TGGGTGCTCACTGCGCCGACCTGCTTCTTGAGACCCGGAGAAAGTCTCGAGAGCTGG 1132
Db 364 TrpValLeuThrAlaAlaHisCysPhePheValThrArgGluLysValLeuGluGlyTrp 383
QY 1133 AAGGTGAACCGGGCGGACACCACTTCGACAGTTCGCTGAGGAGCCCTCATTTGGCCGAG 1192
Db 384 LysValIlyrAlaGlyThrSerAsnLeuHisGlnLeuProGluAlaAlaSerIleAlaGln 403
QY 1193 ATCATCATCAACAGCAATTTACACCGATGAGGAGGACGACTATGACATTCGCCCTCATGCGG 1252
Db 404 IleIleIleAsnSerAsnIlyrThrAspGluGluAspAspIlyrAspIleAlaLeuMetArg 423
QY 1253 CTGTCCAGACCCCTGACCTGTCCGCTCACATCCACCTGCTTGCCTCCCATGCAATGGA 1312
Db 424 LeuSerIlyrProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGly 443
QY 1313 CAGACCTTTAGCTCAATGAGACCTGCTGATCAGAGCTTTGGCAAGACAGGGAACA 1372
Db 444 GluThrPheSerLeuAsnGluThrCysTrpIleThrIlyrPheGlyLysTrpArgGluThr 463
QY 1373 GAGGACAGACATCCCTTCCTCCGGAGGTGACAGGTCATCTCATCGACTTCAAGAAA 1432
Db 464 AspAspIlyrThrSerProPheLeuArgIleValAlaHisLeuIleAspPheLysLys 483
QY 1433 TGCATGACTACTTGTGTCTATGACAGTACCTTACCCCAAGATGATGTGTGTGGGAGAC 1492
Db 484 CysAsnAspIlyrLeuValIlyrAspSerIlyrLeuThrProArgMetCysAlaGlyAsp 503
QY 1493 CTTGTGGGGGAGAGACTCTGTCCAGGAGACAGCGGGGGGCTCTTGTGTGTGACAG 1552
Db 504 LeuArgIlyrGlyArgAspSerCysGlnGlyAspSerGlyIlyrProLeuValCysGlnGln 523
QY 1553 AACAAACCGCTGCTCCCTGGGAGGTGACACGCTGGGGGACAGGCTGTGGCCAGAGAAC 1612
Db 524 AsnAsnAspIlyrIlyrLeuAlaGlyValThrSerIlyrGlyCysGlyGlnAsn 543
QY 1613 AAACCTGATGTGTACCAAGATGACAGAACTTCTCCCTGATTTACAGCAAGATGAG 1672
Db 544 LysProGlyValIlyrThrIlyrValThrGluValLeuProIlyrIleIlyrSerLysMetGln 563

RESULT 10
US-10-428-275-104
; Sequence 104, Application US//10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390

; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450
; SOFTWARE: CuraSequit version 0.1
; SEQ ID NO 104
; LENGTH: 586
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-428-275-104
Alignment Scores:
Pred. No.: 2, 02e-171 Length: 586
Score: 2948.50 Matches: 554
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 0
Query Match: 88.5% Indels: 5
DB: 4 Gaps: 1
US-10-806-370-11 (1-1748) x US-10-428-275-104 (1-586)
QY 11 ATGAGAGGGACAGCCAGGGAATGCATCTTCAGCAAGAACACCTTACGCTGAGACATCT 70
Db 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGCACTCCAGCTGGGAGACACCTCCAGCCGGGCACTCTCCAGCCAGGCACTCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyIlyrGlyAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCAGGCACTCCAGCTGGGAGACACCTCCAGCCGGGCACTCTCCAGCCAGGCACTCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyIlyrGlyAlaSerProAlaGlnAlaSer 60
QY 191 CCAAGTGTACACCTCCAGCCGGGCACTCTCCAGCCGGG-----GCATCT 235
Db 61 ProAlaGlyThrProProGlyIlyrGlyAlaSerProGlyIlyrGlyAlaSerProAlaGlnAlaSer 80
QY 236 CCAGCCAGGCACTCCAGCCGGGCACTCTCCAGCCGGGCTCTGAGCATCATCTTCAGGCTCTCA 295
Db 81 ProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSerLeuSerAspSer 100
QY 296 TCCGGCAGATCATCATCCGCGAGGATCAGCTCGGTGACCACTCCCAACAGAGTGTAC 355
Db 101 SerGlyIlyrSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValIlyr 120
QY 356 CTTGTAGACCAACACCAAGTGGGGGCTGTACCAATCCATCATCTCTCCAGGTACGA 415
Db 121 LeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
QY 416 CCAGCAACCGGGCCACCCAGGAGAGCCGAGTACGACGCTGCCAAGTTCACCTGGCGG 475
Db 141 ProAlaThrArgAlaThrArgIlyrGlySerProGlyIlyrSerLeuProLysPheThrTrpArg 160
QY 476 GAGGGCCAGAGAGAGTACCGGCTCATCGGAGGTGCTCTCTCCATTCGCTGGTGTGT 535
Db 161 GluGlyGlnIlyrGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaValVal 180
QY 536 TCGCTCATCATCTCTTCAGATTCTGGCAGGGCCACACAGGATCAGTACAGAGACAG 595
Db 181 SerLeuIleIleLeuPheGlnPheTrpGlnGlyIlyrIleThrGlyIleArgIlyrLysGlnGln 200
QY 596 AGGAGAGCTGTCCCAAGACACGCTGTTCGCTGACCGGGGTGGTGACTCAAGCTGAAG 655
Db 201 ArgGlnSerCysProLysHisAlaValArgCysAspIlyrAlaValAspCysLysLeuLys 220
QY 656 AGTACAGAGCTGGGCTGCGTAGAGTTTGACTGGAGACAAGTCTGCTTAAATCTACTCT 715
Db 221 SerAspGluLeuGlyCysValArgPheAspTrpAspLysSerLeuLeuLysIleIlyrSer 240
QY 716 GGGTCTCTCCATCATGATGGCTTCCATCTGTAGACAGCACTGGAATGACTCTACTAGAG 775
Db 241 GlySerSerHisGlnTrpLeuProIleCysSerSerAsnTrpAsnAspSerIlyrSerGln 260

Oy 776 AAGACCTGCGACAGCTGGGTTTGGAGAGTGTCAACCGGAACAACGAGGTGGCCACAGG 835
 Db 261 LysThrCysGlnGlnLeuGlyPheGlnSerAlaHisValGlnThrGluValAlaHisValG 280
 Oy 836 GATTTTGCGAACAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGAAAGCTCCAC 895
 Db 281 AspPheAlaIleAsnSerPheSerIleLeuAlaGlyThrAsnSerThrIleGlnGlnSerLeuHis 300
 Oy 896 AGGTGTGAATGACCTTCCCAAGCGGTATATCTCCCTCCAGTGTTCACATGCTCCGACTGAG 955
 Db 301 ArgSerGlnCysProSerGlnAlaGlyThrIleSerLeuGlnCysSerHisCysGlyLeuAlaG 320
 Oy 956 GGCATGACCGGGCGGATGTGTGGAGAGGGCGGTGCGCTCCGATACCAAGTGGCCCTGGCAA 1015
 Db 321 AlaIleThrGlyAlaGlyIleValGlyAlaLeuAlaSerAspSerLysTrpProTrpGln 340
 Oy 1016 GTGAGTCTGCACCTTCGAGCACCAACCCATCTGTGGAGGACGCTCATTTGACGCGCAATGG 1075
 Db 341 ValSerLeuHisAspPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrp 360
 Oy 1076 GTGCTCACTGCGCGCCCACTGCTTCTTCGTGACCCCGGAGAAAGTCTTGAGAGCTGGAAG 1135
 Db 361 ValLeuThrAlaAlaHisCysAspPheValThrArgGluLysValLeuGlnGlyTrpLys 380
 Oy 1136 GTGTACCGGGGACACCGAACCTTGACCACTGCTGCTGAGGACAGCTTCATTTCCGAAATC 1195
 Db 381 ValIleThrAlaGlyThrSerAsnLeuHisGlnLeuProGluAlaAlaSerIleAlaGlnIle 400
 Oy 1196 ATCATCAACAGCAATTAACCGGATGAGGAGGACGACTGATCATCGACCTTCATCGAGCTG 1255
 Db 401 IleIleAsnSerAsnThrThrAspGluGlnAspIleAspIleAlaAlaMetArgLeu 420
 Oy 1256 TCCAAAGCCCTGACCCCTGTCGCGCTCACATCCACCTGCTTCGCTCCCATCGATGACAG 1315
 Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGln 440
 Oy 1316 ACCTTTAGCTTCATGAGACCTGCTGATCAACAGGCTTTGGCAAGACCGAGGACAGAT 1375
 Db 441 ThrPheSerLeuAsnGlnThrCysThrTrpIleThrGlyPheGlyLysThrArgGlnThrAsp 460
 Oy 1376 GACAAGACATCCCCCTCTCCCGGAGGTGGAGGCAATCTCATCGAATTCAGAAATGC 1435
 Db 461 AspLysThrSerProPheLeuArgGluValGlnValAsnLeuIleAspPheLysLysCys 480
 Oy 1436 AATGACTACTTGCTATGACAGATTACCTTACCCCAAGATGATGTGCTGCGGAACTT 1495
 Db 481 AsnAspIleLeuValIleAspSerIleuThrProArgMetMetCysAlaGlyAspLeu 500
 Oy 1496 CGTGGGGGACAGACCTCTGCGCAGGAGACAGCGGGGGCTCTTGTCGTGACAGAAC 1555
 Db 501 ArgGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsn 520
 Oy 1556 AACCGCTGTCACCTGGACAGGTGTCAACCAAGCTGGGGACAGGCTGTGGCCAGAAACAA 1615
 Db 521 AsnArgTrpIleLeuAlaGlyValIleThrSerTrpGlyThrGlyCysGlyGlnArgAsnLys 540
 Oy 1616 CCGGTGTGTACACCAAGTACAGAAAGTTCTTCCCGATGATTTACAGACAGATGAG 1672
 Db 541 ProGlyValIleThrLysValThrGluValLeuProTrpIleIleIleSerLysMetGln 559
 RESULT 11
 US-10-428-275-124
 : Sequence 124, Application US/10428275
 : Publication No. US20040067505A1
 : GENERAL INFORMATION:
 : APPLICANT: Alvarez et al.
 : TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
 : FILE REFERENCE: 21402-585
 : CURRENT APPLICATION NUMBER: US/10/428, 275
 : CURRENT FILING DATE: 2003-05-01
 : PRIOR APPLICATION NUMBER: 09/966545
 : PRIOR FILING DATE: 2001-09-26
 : PRIOR APPLICATION NUMBER: 09/544511

```

: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: 60/128514
: PRIOR FILING DATE: 1999-04-09
: PRIOR APPLICATION NUMBER: 09/569269
: PRIOR FILING DATE: 2000-05-11
: PRIOR APPLICATION NUMBER: 60/134315
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 09/619252
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/789390
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185548
: PRIOR FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 450
: SOFTWARE: Curaseq1set version 0.1
: SEQ ID NO 124
: LENGTH: 586
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-428-275-124

Alignment Scores:
Pred. No.: 2, 02e-171
Score: 2948.50
Percent Similarity: 99.1%
Best Local Similarity: 99.1%
Query Match: 88.5%
DB: 4 Gaps: 1

US-10-806-370-11 (1-1748) x US-10-428-275-124 (1-586)

QY 11 ATGAGAGGAGCAGCCAGCGAATCATCTCCAGCAGAACACCTTCACTGAGCATCT 70
Db 1 MetGlunArgAspSerHisIleGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20

QY 71 CCAAGCCAGGCATCTCCAGCTGGAGACCTCCAGGCCGGGATCTCCAGCCCGAGCATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40

QY 131 CCAAGCCAGGCATCTCCAGCTGGAGACACTCCGGGCGGGGATCTCCAGCCAGGCATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60

QY 191 CCAAGCTGTACACTCCAGCCAGGCCAGCATCTCCAGGCCG-----GCATCT 235
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80

QY 236 CCAAGCCAGGCATCTCCAGCCCGGGGATCTCCGGCTCTGGCATCACTTCCAGGTCCTCA 295
Db 81 ProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSer 100

QY 296 TCCGGGAGGCATCACTCCGCCCGAGGTCAGCTCCGGGAGCAACCTCCCAACAGAGGTAC 355
Db 101 SerGlyArgSerSerSerAlaArgSerAlaSerAlaThrThrSerProThrArgValTyr 120

QY 356 CTGTGTAGAGCAACACCAAGTGGGGGCTGATCCATCCGATCATCTCTCCAGGTCAGCA 415
Db 121 LeuValAlaArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140

QY 416 CCAAGCAACCAAGGCGCCACAGGAGAGCCAGGATCAGACCTGCCAAGTTCACTGGCGG 475
Db 141 ProAlaThrArgAlaThrArgIleArgIleSerProGlyThrThrSerLeuProLysPheThrTyrArg 160

QY 476 GAGGGGCAAGAGAGATCAAGCTCAATCCGGGTCGGGTCCTCCATCATTCGCTGGAGTT 535
Db 161 GluGlyGlnIlePheIleuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValVal 180

QY 536 TCAGCTATCATCTCTTCAGTTCTGGAGGGCCACAGAGGATCAGGTACAAAGAGACAG 595
Db 181 SerLeuIleIleLeuPheGlnPheThrArgIleThrGlyIleArgTyrLysGlnGln 2000

QY 596 AGGAGAGAGCTGTCCAAAGACGCTGTTCGCTGTGACGGGGTGTGTGACTGCACAGCTGAG 655

```

```

Db      201 ArgGluserCysProlysh:laValaValArgCysAspGlyValaValaPserCysLysLeuLys 220
QY      656 AGTGAAGAGCTGGGCTCCGCTGAGGTTGACTGGGACAAGTCTCGCTTAATATCTCT 715
        |||
Db      221 SerAspGluLeuGlyCysValaArgPheAspTrpAspLysSerLeuLysLysIleTyrSer 240
QY      716 GGGTCCCGCATCAGTGGCTTCCCATCTGTAGCAGCAAGTGAATGACTCTACTCAGAG 775
        |||
Db      241 GlySerSerHisGlnTrpLeuProIleCysSerSerLeuTrpAspMetPheSerGlu 260
QY      776 AAGACCTCCAGCAGACTGGGTTTCGAGAGTCTCACCGCAACCGAGTTCGCCACAG 835
        |||
Db      261 LysThrCysGlnGlnLeuGlyPheGluSerAlaHisArgThrThrGluValAlaHisArg 280
QY      836 GATTTCGCCAAGCTTCTCATATTGAGATACACTCCACATCCAGAAAGCTTCAC 895
        |||
Db      281 AspPheAlaAsnSerPheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuHis 300
QY      896 AGGCTGAATGGCTTCCCGCGGATATATCTCCCTCCAGTTCSCACTGGCGAGCTGAG 955
        |||
Db      301 ArgSerGluCysProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
QY      956 GCCATGACCGGCGGATCTGTGGAGGCGCGCTGGCTCGGATAGCAAGTGGCTTGCAA 1015
        |||
Db      321 AlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTrpProTrpGln 340
QY      1016 GTGAGTGTGCACTTGGGACCAACCCACATCTGTGGAGGCAAGCTCATTTAGCCCACTG 1075
        |||
Db      341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrp 360
QY      1076 GTGTCACTGCGCGGCACTGTTCTTGTCGACCGGGAGAAGGCTGGAGGCTGGAG 1135
        |||
Db      361 ValLeuThrAlaAlaHisCysPhePheValThrArgGluValaLysGluGlyTrpLys 380
QY      1136 GTGTACCGGGGCAACCAACCTGCAACAGTGGCTGAGGCAAGCTCATTTGCCAGATC 1195
        |||
Db      381 ValTyrAlaGlyThrSerLeuLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIle 400
QY      1196 ATCATCAACAGCAATTAACCCGATGAGAGAGAGCACTATGACATGCCCTTCAGCGCTG 1255
        |||
Db      401 IleIleAsnSerAsnTyrThrAspGluGluAspAspTyrAspIleAlaLeuMetArgLeu 420
QY      1256 TCAGAGCCCTGACCCCTGTCGCTGACATCAACCTGCTGCTGCCCAAGCAAGAGCAG 1315
        |||
Db      421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGln 440
QY      1316 ACCTTAGCTCAATGAGACCTGCTGGATCAAGAGCTTTGGCAAGACAGAGAGACAGAT 1375
        |||
Db      441 ThrPheSerLeuAsnGlnThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAsp 460
QY      1376 GACAAAGACATCCCTTCTCCCGGAGGTGCAAGTCAATCTCATGCACTTCAGAAATGC 1435
        |||
Db      461 AspLysThrSerProPheLeuArgGluValGlnValaAsnLeuIleAspPheLysLysCys 480
QY      1436 AATGACACTTGTGCTTGAAGACTTACCTTAACCCCAAGATGATGTGTGCTGGGAGACTT 1495
        |||
Db      481 AsnAspLysTrpLeuValTyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeu 500
QY      1496 CGTGGGGGAGAGACTCTGTCAGAGGAGACAGCGGGGGCTTGTGCTGTGACAGAAC 1555
        |||
Db      501 ArgGlyGlyLysArgAspSerCysGlnGlyAspSerGlyLysProLeuValCysGlnGlnAsn 520
QY      1556 AACCGCTGTACTGTGAGGTGTGACAGCTGGGGCAAGGCTGTGGCCAGAGAAACAA 1615
        |||
Db      521 AsnArgTrpTyrLeuAlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAspLys 540
QY      1616 CTGGGTGTGACCAAGAGAGAGAGAGTCTTCCCTGGATTTACAGACAGATGAG 1672
        |||
Db      541 ProGlyValTyrThrLysValThrGluValLeuProTrpIleTyrSerLysMetGln 589

```

```

: Publication No. US20040067505A1
: GENERAL INFORMATION:
: APPLICANT: Alvaraz et al.
: TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
: FILE REFERENCE: 21402-585
: CURRENT APPLICATION NUMBER: US/10/428,275
: CURRENT FILING DATE: 2003-05-01
: PRIOR APPLICATION NUMBER: 09/966545
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 09/544511
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: 60/128514
: PRIOR FILING DATE: 1999-04-09
: PRIOR APPLICATION NUMBER: 09/569269
: PRIOR FILING DATE: 2000-05-11
: PRIOR APPLICATION NUMBER: 60/134315
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 09/619252
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/789390
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185548
: PRIOR FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 450
: SOFTWARE: CuroSeqList version 0.1
: SEQ ID NO 142
: LENGTH: 586
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-428-275-142

```

```

Alignment Scores:
Pred. No.: 4.07e-171 Length: 586
Score: 2943.50 Matches: 553
Percent Similarity: 98.9% Conservative: 0
Best Local Similarity: 98.9% Mismatches: 1
Query Match: 88.3% Indels: 5
DB: 4 Gaps: 1

```

US-10-806-370-11 (1-1748) x US-10-428-275-142 (1-586)

```

QY      11 ATGAGAGGAGCAGCAGCGGAATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 70
        |||
Db      1 MetGluArgAspSerHisGlyAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY      71 CCAAGCCAGGCATCTCCAGCTGGGACACTCCAGGCGGGCATCTCCAGCCAGGCATCT 130
        |||
Db      21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY      131 CCAAGCCAGGCATCTCCAGCTGGGACACTCCAGGCGGGCATCTCCAGCCAGGCATCT 190
        |||
Db      41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY      191 CCAAGCTGTACCTCCAGGCGGGCATCTCCAGGCGG-----GCATCT 235
        |||
Db      61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY      236 CCAAGCCAGGCATCTCCAGGCGGGCATCTCCAGCTGGGCATCACTTCAGGCTCTCA 295
        |||
Db      81 ProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSer 100
QY      296 TCCGGCAGGTCATCATCTCCGCGCAGGTCAGCTCGGTGCAACCTCCCAACCAAGGTGAC 355
        |||
Db      101 SerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
QY      356 CTGTGTAGACAACACAGAGTGGGGGTGTACCATCCGATCATCTCTGCGCAGGTACGA 415
        |||
Db      121 LeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
QY      416 CCAAGCAACAGGGCCACAGGAGAGAGCCCAAGGTACAGGCTGCCCAAGTTCACTGGCGG 475
        |||
Db      141 ProIleThrArgAlaThrArgGlnSerProGlyThrSerLeuProLysPheThrTrpArg 160

```


Db 101 SerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
 QY 356 CTGTTGAGAAACACAGTGGGGCTGTACCCATCCGATATCTCTGCGAGGTGACA 415
 Db 121 LeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
 QY 416 CCAGCAACAGGGGCGACAGGAGAGAGCCAGGTACAGACCTGCGCCAGTTACCTGGCGG 475
 Db 141 ProAlaThrArgAlaThrArgGlySerProGlyThrSerLeuProLysPheThrTyrArg 160
 QY 476 GAGGCGCAAGAGACGTACCGCTGCGGGTGGCTCTCTCATTTGCGCTGGTGGT 535
 Db 161 GlyGlyGlyLeuValGlyLeuProLeuIleGlyValLeuLeuLeuIleAlaLeuValVal 180
 QY 536 TCGGTATCATCTCTTCCAGTTCTGGGAGGGCCACACAGGATCAGGTACAGAGAGAG 595
 Db 181 SerLeuIleIleLeuPheGlnPheThrGlnGlyIleThrGlyIleArgGlyLeuGln 200
 QY 596 AGGAGAGAGCTGTCCCAAGACAGCTGTTCGCTGTACGGGGTGGTGGACTGCAAGTGA 655
 Db 201 ArgGlySerCysProLysPheValAlaValArgCysArgGlyValValArgCysLeuLys 220
 QY 656 AGTACAGAGCTGGGCTGCGTGAAGTTTACTGGAGCAAGTCTCTTAAATCTACTCT 715
 Db 221 SerArgGlyLeuGlyCysValArgPheAspThrAspLysSerLeuLeuLysIleTyrSer 240
 QY 716 GGGTCTCCCATCATGTGGCTTCCATCTGTAGCAAGCACTGGAAATGACTCTTACTCAGAG 775
 Db 241 GlySerSerHisGlyThrLeuProIleCysSerSerHisThrAsnAspSerTyrSerGlu 260
 QY 776 AAGACCTGCCACAGCTGGGTTTGGAGAGTCTCACCGGAAACCGAGGTTGGCCACAG 835
 Db 261 LysThrCysProGlnLeuGlyPheGlySerAlaHisValGlyThrGlnValAlaHisVal 280
 QY 836 GATTTCGCAACAGCTTCTCAATCTTGAGATACATCCACATCCAGCAAGAAAGCTCCAC 895
 Db 281 AspPheAlaAsnSerPheSerIleLeuArgGlyThrAsnSerThrIleGlnGlnSerLeuHis 300
 QY 896 AGGTCTGAATGCTTCCACAGCGGTATATCTCCCTCAAGTTTCCACCTGGAGCTGAAG 955
 Db 301 ArgSerGlyCysProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
 QY 956 GGCATGACCGGGGCGATCGTGGGAGGGGCGCTGGCTCGATACAGTGGCTTTGGCAA 1015
 Db 321 AlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTyrProIlePheGln 340
 QY 1016 GTGAGTCTGCACTTGGCAGCCACCATCATCTGTGGAGGCAAGCTCATTTGACGCCAGTGG 1075
 Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnThr 360
 QY 1076 GTGCTCACTGCGCGGCACTGCTTTTGTGTGACCGGGAGAAAGTCTGGAGGGCTGAAAG 1135
 Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgGlyValLeuGlyGlyTyrLys 380
 QY 1136 GTGAGCGGGGACACAGCAACCTGACAGTGGCTGAGGAGGCTCATTTGCGCGAATC 1195
 Db 381 ValTyrAlaGlyThrSerAsnLeuHisGlnLeuProGlyAlaAlaSerIleAlaGlnIle 400
 QY 1196 ATCATCAACAGCAATTACACCGATGAGAGAGAGCACTATGACATCGCTCATCGGCTG 1255
 Db 401 IleLeuAsnSerAsnTyrThrAspGlyGlyAspArgTyrAspIleAlaLeuMetArgLeu 420
 QY 1256 TCCAGGCGCTGACCTGTGCGCTCAGATCCACCTGCTTGGCTCCCATCGATGAGACAG 1315
 Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProLysCysLeuProMetHisGlyGln 440
 QY 1316 ACCCTTGACCTGATGAGACCTGTGATGACAGAGCTTTGGCAGAGCAGGAGAGACAT 1375
 Db 441 ThrPheSerLeuAsnGlnThrCysTyrPheThrGlyPheGlyLysThrArgGlyThrAsp 460
 QY 1376 GACAAGACATCCCTCTCTCCGGAGAGTGCAGGTCAATCTCATGCACTTCCAGAAATGC 1435
 Db 461 AspArgThrSerProPheLeuArgGlyValGlnValAsnLeuIleAspPheLysLysCys 480

QY 1436 AATGACTACTTGGCTATGACAGTTACTTACCCCAAGAAATGATGTCGGGGACCTT 1495
 Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeu 500
 QY 1496 CTGGGGGCGAGAGACTCTCTGCGAGGAGACAGCGGGGGCTCTTGTCTGTGACAGAAC 1555
 Db 501 ArgGlyGlyArgAspSerCysGlnGlyAspSerIleGlyProLeuValCysGlyGlnAsn 520
 QY 1556 AACCGCTGGTACTGCGAGGTGTACACAGCTGGGGCACAGGCTGTGCGCAGAGAAACAA 1615
 Db 521 AsnArgTyrTyrLeuAlaGlyValThrSerThrGlyCysGlyGlnArgAsnLys 540
 QY 1616 CTTGGTGTGACACCAAGTGACAGAGTCTTCCCGAATTACAGAGAGATGAG 1672
 Db 541 ProGlyValTyrThrLysValThrGlnValLeuProIleTyrSerLysMetGlu 559
 RESULT 14
 US-09-888-615-104
 : Sequence 104, Application US/09888615
 : Patent No. US20020064856A1
 : GENERAL INFORMATION:
 : APPLICANT: PLOMMAN, GREGORY
 : APPLICANT: WHYTE, DAVID
 : APPLICANT: CAENEDEBEL, SEAN
 : APPLICANT: CHARNOZAK, GLEN
 : APPLICANT: MANNING, GERARD
 : APPLICANT: SUDARSANAM, SUCHA
 : TITLE OF INVENTION: NOVEL PROTEASES
 : FILE REFERENCE: 038602/1214
 : CURRENT APPLICATION NUMBER: US/09/888,615
 : PRIOR FILING DATE: 2001-06-26
 : PRIOR APPLICATION NUMBER: 60/214,047
 : NUMBER OF SEQ ID NOS: 150
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 104
 : LENGTH: 537
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-888-615-104
 Alignment Scores:
 Pred. No.: 7,07e-161 Length: 537
 Score: 2775.50 Matches: 526
 Percent Similarity: 92.0% Conservative: 0
 Best Local Similarity: 92.0% Mismatches: 1
 Query Match: 83.3% Indels: 45
 DB: 3 Gaps: 2
 US-10-806-370-11 (1-1748) x US-09-888-615-104 (1-537)
 QY 11 ATGAGAGAGGACAGCCAGGGAATGATCTTCAGCAAGAAACCTTGAGTGAAGATCT 70
 Db 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CAGGCCAGGATCTCAGCTGGGACACTTCAGGCGGGGATCTTCAAGCCAGGATCT 130
 Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CAGGCCAGGATCTCAGCTGGGACACTTCAGGCGGGGATCTTCAAGCCAGGATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CAGCTGTGACACTCCAGGCGGGGATCTTCAGGCGG----- 229
 Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 230 -----GCATCTCAGGCCAGGCACTTCAGCCCGGCGATCTCCGGCTTGGCATCA 280
 Db 81 ProAlaGlnAlaSerProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSer 100
 QY 281 CTTTCCAGTCTTCATCCGGCAGGTATCATCCGCCAGTCCAGCTCGGTGACAACTCC 340


```

Db      |||...|||
101 LeuSerArgSerSerGlyArgSerSerLeuAlaArgSerAlaSerValThrThrSer 120
Qy      |||...|||
341 CCAACGAGGTGATCTTTAGAGCAACACAGTGGGGGCTGTAACCATCCGATCT 400
Db      |||...|||
121 ProThrArgValIleuValArgAlaThrProValGlyAlaValProIleArgSerSer 140
Qy      |||...|||
401 CTGCGCAGGTGAGCAGCAGCAACGAGGCCACGAGGAGAGCCGAGTACAGCTGCC 460
Db      |||...|||
141 ProIleArgSerAlaProAlaThrArgAlaThrArgIleuSerPro----- 155
Qy      |||...|||
461 AAGTTCACTGGCGGAGGAGCCAGAGACGCTACCGCTCAGGTGGCTGCTCTCTC 520
Db      |||...|||
155 ----- 155
Qy      |||...|||
521 ATTGCCTGGTGGTTTGGCTGATCATCTCTTCAGATTTCGGAGGGGCACAGGGATC 580
Db      |||...|||
156 -----ValGlnPheTrpGlnGlyHisThrGlyIle 165
Qy      |||...|||
581 AGGTACAAGGAGCAGAGGAGAGCTGTCCAGACGCTGTTCCGCTGTGACGGGGTGGT 640
Db      |||...|||
166 ArgTyrIleGlnGlnArgIleuSerCysProIleHisAlaValArgCysAspGlyValVal 185
Qy      |||...|||
641 GACTGCAAGCTGAAGATGACAGCTGGCTGGCTGAGTTGACTGGAGCAAGTCTCTG 700
Db      |||...|||
186 AspCysIleuSerIleuSerAspGlnLeuGlyCysValArgPheAspTrpAspIleu 205
Qy      |||...|||
701 CTTAAATCTACTGTGGATCTCCCATCAGAGGCTCCCATCTGTAGAGCAACTGGAA 760
Db      |||...|||
206 LeuIleIleTyrSerGlySerSerHisGlnTrpLeuProIleCysSerSerHisTrpShn 225
Qy      |||...|||
761 GACTCTACTCAGAGAAGACCTGCCAGACGCTGGGTTTGAAGTGTCTCAGCGGACA 820
Db      |||...|||
226 AspSerTyrSerGlnIleuSerHisCysGlnGlnIleuGlyPheGlnSerAlaHisArgThrThr 245
Qy      |||...|||
821 GAGGTTCCCAACAGGATTTTGGCCACAGCTTTCATCTTGAATACACTCCACCATC 880
Db      |||...|||
246 GluValAlaHisArgAspPheAlaAsnSerPheSerIleuArgTyrAsnSerThrIle 265
Qy      |||...|||
881 CAGGAAGACCTCCACAGATCTGAATGCCCTTCCAGAGGATATCTCCCTCAGTGTCC 940
Db      |||...|||
266 GlnIleuSerLeuHisArgSerGlnCysProSerGlnArgTyrIleSerLeuGlnCysSer 285
Qy      |||...|||
941 CACTGCGAGCTGAGGCGCATGACCGGCGAGATCTGGGAGGCGGCTGCGCTCGATAGC 1000
Db      |||...|||
286 HisCysGlyLeuValArgAlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSer 305
Qy      |||...|||
1001 AAGTGGCTTTGGCAAGTGAAGTCTGCACTTGGGACACCCACATCTGTGAGGACGCTC 1060
Db      |||...|||
306 LysTrpProTrpGlnValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeu 325
Qy      |||...|||
1061 ATTGACGCCAGTGGAGTGTCTACGCGGCCCATGCTTCTTGGAGCCCGGAGAAAGGTC 1120
Db      |||...|||
326 IleAspIleGlnTrpValIleuThrAlaHisIleCysPhePheValIleArgGlyValVal 345
Qy      |||...|||
1121 CTGAGAGGCTGGAAGGTGTAGCGCGGCAACGACACCTGCAACAGTTGCCCTGAGGACCC 1180
Db      |||...|||
346 LeuGlnGlyTrpIleValIleGlyThrSerAsnLeuHisGlnLeuProGlnAlaAla 365
Qy      |||...|||
1181 TTCATTCCCGAGATCATCATCAACGCAATTAACCCGATGAGGAGACGACTATGACATC 1240
Db      |||...|||
366 SerIleIleGlnIleIleIleAsnSerAsnIleThrAspGlnIleuAspAspIle 385
Qy      |||...|||
1241 GCCCTCATGGCGGCTGTCCAGAGCCCTGACCTGCGGCTGACATCCACCCGCTGTGCTC 1300
Db      |||...|||
386 AlaLeuMetArgLeuSerIleProLeuThrIleuSerAlaHisIleIleProAlaCysIleu 405
Qy      |||...|||
1301 CCATGATGACAGACCTTAAAGCTCAATGACACTGTGGATCAGAGGCTTTGGCAAG 1360
Db      |||...|||
406 ProMetHisGlyGlnThrIlePheSerIleuAsnGlnIleuThrCysTrpIleThrGlyPheGlyLys 425
Qy      |||...|||
1361 ACCAGGAGACAGATGACAGACATCCCTCTCTCCGAGGAGTGCAGGTCAATCTCATC 1420

```

```

Db      |||...|||
426 ThrArgIleuThrAspAspLeuThrSerProPheLeuArgIleuGlnValAsnLeuIle 445
Qy      |||...|||
1421 GACTTCAAGAAATGCAATGACTACTTGATCTATGACAGATTACTTACCCCAAGGATGATG 1480
Db      |||...|||
446 AspPheIleuSerCysAsnAspTrpLeuValIleArgAspSerTyrLeuThrProArgMetMet 465
Qy      |||...|||
1481 TTGCTGGGGACCTTGTGGGGGAGAGACTCTCTCCAGGAGAGACAGCGGGGGCTCTT 1540
Db      |||...|||
466 CysAlaGlyAspLeuArgIleGlyArgAspSerCysGlnGlyAspSerGlyIleProLeu 485
Qy      |||...|||
1541 GTCTGTGACAGAACCAACCCCTGTGTTACTGTCAGAGTGTCCACCACTGGGCGACAGGCT 1600
Db      |||...|||
486 ValCysGlnGlnIleuAsnAsnArgTrpTrpLeuValGlyAlaThrSerTrpGlyThrGlyCys 505
Qy      |||...|||
1601 GGCCAGAGAAACAAACCTGGTGTATACACCAAGTACAGAAAGTTCTTCCCTGGATTAC 1660
Db      |||...|||
506 GlyIleArgAsnIleProGlyValIleThrIleValIleThrGlnValIleuProTrpIleTyr 525
Qy      |||...|||
1661 AGCAAGATGAGAGAGCGAGGTGCGATTGAGAAATCC 1696
Db      |||...|||
526 SerIleMetGlnSerGlnValArgPheArgIleSer 537

RESULT 15
US-10-428-275-128
: Sequence 128, Application US/10428275
: Publication No. US20040067505A1
: GENERAL INFORMATION:
: APPLICANT: Alvarez et al.
: TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
: FILE REFERENCE: 21402-588
: CURRENT APPLICATION NUMBER: US/10/428,275
: PRIOR FILING DATE: 2003-05-01
: PRIOR APPLICATION NUMBER: 09/966545
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 09/544511
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: 60/128514
: PRIOR FILING DATE: 1999-04-09
: PRIOR APPLICATION NUMBER: 09/569269
: PRIOR FILING DATE: 2000-05-11
: PRIOR APPLICATION NUMBER: 60/134315
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 09/619252
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/789390
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185548
: PRIOR FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 450
: SOFTWARE: Curoseqdist version 0.1
: SEQ ID NO 128
: LENGTH: 537
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-428-275-128

Alignment Scores:
Pred. No.: 7,07e-161 Length: 537
Score: 2775,50 Matches: 526
Percent Similarity: 92,0% Conservative: 0
Best Local Similarity: 92,0% Mismatches: 1
Query Match: 83,3% Indels: 45
DB: 4 Gaps: 2

US-10-806-370-11 (1-1748) x US-10-428-275-128 (1-537)
Qy      |||...|||
11 ATGAGAGGAGCAGACCGAGGAATGCATCTCCAGCAAGAACACCTTCACTGAGCATCT 70
Db      |||...|||
1 MetGlnArgAspSerHisGlyAsnAlaSerProIleArgThrProSerAlaGlyAlaSer 20
Qy      |||...|||
71 CCGAGCCAGGACATCTCAGCTGGGAGCACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 130
Db      |||...|||
21 ProIleGlnAlaSerProIleAlaGlyThrProGlyIleArgAlaSerProIleGlnAlaSer 40

```


QY 131 CCAAGCCAGGATCTCCAGCTGGAGACCTCCGGGCGGGGATCTCCAGCCAGGATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyThrAlaSerProAlaGlnAlaSer 60
 QY 191 CCAAGCTGGTACCTCCAGGCGGGGATCTCCAGCCCGG----- 229
 Db 61 ProAlaGlyThrProProGlyThrAlaSerProGlyThrAlaSerProAlaGlnAlaSer 80
 QY 230 -----GCATCTCAAGCCAGGATCTTCAGCCCGGGGATCTCCGGCTCTGGCATCA 280
 Db 81 ProAlaGlnAlaSerProAlaGlnAlaSerProAlaGlnAlaSerProAlaGlnAlaSer 100
 QY 281 CTTTCCAGGCTCCATCCGGAGGCTCATCCGGCCAGGCTGAGCTGGGTGACAACTCC 340
 Db 101 LeuSerArgSerSerSerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSer 120
 QY 341 CCAACCAAGTGTACTTGTAGAGCAACCAAGTGGGGGCTGTACCATCCATCCATCATCT 400
 Db 121 ProThrArgValThrLeuValArgAlaThrProValGlyAlaValProIleArgSerSer 140
 QY 401 CTTGCCAGGTACAGCAACCAAGGCGCAACAGGAGAGCCCAAGGTACGAGCTGCC 460
 Db 141 ProAlaArgSerAlaProAlaThrArgAlaThrArgGlySerPro----- 155
 QY 461 AAGTTCACCTGGGGGAGGGCCAGAACAGCTACCGCTCATCGGGGTGCGTCTCTCTC 520
 Db 155 ----- 155
 QY 521 ATTGCCGTGGTGTTCGCTCATCATCTCTTCAGTTCCTGGCAGGGCCACACAGGATC 580
 Db 156 -----ValGlnHeTrpGlnGlyHisThrGlyIle 165
 QY 581 AGGTACAGAGAGCAGAGGAGAGCTGTCCAGACCGCTGTCCCTGTGACGGGCTGTG 640
 Db 166 ArgGlyTrpGlnGlnArgGlySerCysProIleHisAlaValArgCysAspGlyValVal 185
 QY 641 GACTGCAAGCTGAGATGACAGCTGGCGCTGGCGGTGAGTTCGAGAACTCTCTG 700
 Db 186 AspCysLeuLeuLeuSerSerSerGlnLeuGlyCysValArgPheAspTrpAspLeuSerLeu 205
 QY 701 CTTAAATCTACTCTGGGCTCTCCCATCAGTGGCTTCCATCTGTAGCAGCACTGGAAT 760
 Db 206 LeuLeuIleLeuTrpSerGlySerSerHisGlnTrpLeuProIleCysSerSerAsnTrpAsn 225
 QY 761 GACTCTACTCAGAGAAAGCTGCGCAGAGCTGGGTTTCGAGAGTCTCACGGACAAC 820
 Db 226 AspSerTrpSerGlnLeuLeuThrCysGlnGlnLeuGlyPheGlnSerAlaHisArgThrThr 245
 QY 821 GAGGTTCGCCACAGGATTTTGGCAACAGCTTTCATCTTGAGATACAACTCCACCATC 880
 Db 246 GluValAlaHisArgAspPheAlaAsnSerPheSerIleLeuArgGlyTrpAsnSerThrIle 265
 QY 881 CAGGAAAGCCCTCCAGAGTCTGAATGCCCTTCCAGCGGTATATCTCCCTCAGTGTCC 940
 Db 266 GlnGlnSerLeuHisArgSerGlnCysProSerGlnArgGlyTrpIleSerLeuGlnCysSer 285
 QY 941 CACTGCGAGTAGAGGACCATGACCGGGCGAGTCGTGGAGGGGCGCTGGCTCGAATAGC 1000
 Db 286 HisCysGlyLeuValArgAlaMetThrGlyArgGlyLeuValGlyAlaLeuAlaSerAspSer 305
 QY 1001 AAGTGGCTTGGCAAGTGAAGTCTGAGCAACCCCACTGTGAGGAGGACGCTC 1060
 Db 306 LysTrpProTrpGlnValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeu 325
 QY 1061 ATTGACGCCAGTGGGTGCTCACTGCGCGCCACTGCTTCTGTGACCCGGGAGAGGATC 1120
 Db 326 IleAspAlaGlnTrpValLeuThrAlaAlaHisCysPhePheValThrArgGlyLeuVal 345
 QY 1121 CTGAGAGGCTGGAAGGTGTACCGGGGCAACAGCAACTGCAACGTTGCTGAGGACGCC 1180
 Db 346 LeuGlnGlyTrpLeuValThrAlaGlyThrSerAsnLeuHisGlnLeuProGlnAlaAla 365

QY 1181 TCCATTGCCAGATCATCATCAACAGCAATTACACCGATGAGAGAGCATATGATC 1240
 Db 366 SerIleAlaIleIleIleIleAsnSerAsnThrThrAspGlnGlnAspAspTrpAspIle 385
 QY 1241 GCCCTCATGCGGCTGTGCCAAGCCCTGACCCCTGTCCCTCATCATCCCTGCTGCTC 1300
 Db 386 AlaLeuMetArgLeuSerTrpSerProLeuThrLeuSerAlaHisIleHisProAlaCysLeu 405
 QY 1301 CCCATGATGAGACAGACCTTACCTCAATGAGACCCGTGCGATTCACAGGCTTTGGCAG 1360
 Db 406 ProMetHisGlyGlnThrPheSerLeuAsnGlnThrCysTrpIleThrGlyPheGlyLys 425
 QY 1361 ACCAGGAGACAGATGACACAGATCCCTTCCTCCGAGAGGTGACGTCATCTCATC 1420
 Db 426 ThrArgGlnThrAspAspLysThrSerProPheLeuArgGlnValGlnValAsnLeuIle 445
 QY 1421 GACTTCAGAAATGCAATGACTACTGTGTATGACAGTTACTTACCCCAAGGATGATG 1480
 Db 446 AspPheLysLysCysAsnAspTrpLeuValThrAspSerTrpLeuThrProArgMetMet 465
 QY 1481 TGTCTGGGAGCTTCTGTGGGGGCAAGACTCTCCGCAAGGAGACACCGGGGCTCTT 1540
 Db 466 CysAlaGlyAspLeuArgGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeu 485
 QY 1541 GTCTGTGAGCAGAAACAACCGCTGGTACTGACAGGTGTCAACAGCTGGGGCAGAGCTGT 1600
 Db 486 ValCysGlnGlnAsnAsnArgTrpTrpLeuAlaGlyAlaThrSerTrpGlyThrGlyCys 505
 QY 1601 GGCCAGAGAAACAACCTGTGTGTGTACACCAAGTGCAGAACTTCTCCCTGGAATTTAC 1660
 Db 506 GlyGlnArgAsnLysProGlyValThrThrLeuValThrGlnValLeuProTrpIleTrp 525
 QY 1661 AGCAAGATGAGAGCGGAGGTGCGATTCAGAAATCC 1696
 Db 526 SerLysMetGlnSerGlnValArgPheArgLysSer 537

Search completed: September 16, 2006, 02:41:55
 Job time : 327.5 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: September 16, 2006, 02:42:09 ; Search time 10 Seconds
(without alignments)
3704.748 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 3333
Sequence: 1 cccagagaccatgcagagagga.....ggctgctgcgactcagagaa 1748

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 256596 segs, 70647373 residues
Total number of hits satisfying chosen parameters: 513192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xld
-O=/abs/ABSSWEB.spool/US10806370/runat.15092006.105651.10812/app_query.fasta.1
-DB=Published.Applications_AA_New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPT=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10806370 @CGN 1.1 68 @runat.15092006.105651.10812 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -IONLOG -DEV TIMEOUT=120
-WARM TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New:
1: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/prodata/1/pubppaa/PGT_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------------------------|
| 1 | 729.5 | 21.9 | 453 | 6 | US-10-136-749-64 Sequence 64, Appl1 |
| 2 | 726 | 21.8 | 454 | 7 | US-11-400-825-2 Sequence 2, Appl1 |
| 3 | 694 | 20.8 | 432 | 6 | US-10-196-749-330 Sequence 330, App |
| 4 | 694 | 20.8 | 432 | 7 | US-11-101-316-112 Sequence 112, App |
| 5 | 694 | 20.8 | 432 | 7 | US-11-376-673-112 Sequence 112, App |
| 6 | 608.5 | 18.3 | 416 | 7 | US-11-208-257-8 Sequence 8, Appl1 |
| 7 | 543.5 | 16.3 | 305 | 7 | US-11-293-697-3002 Sequence 3002, Ap |
| 8 | 532.5 | 16.0 | 853 | 6 | US-10-530-187-242 Sequence 242, Appl1 |
| 9 | 532.5 | 16.0 | 855 | 7 | US-11-254-185-2 Sequence 2, Appl1 |

| | | | | | | |
|----|-------|------|-----|---|-------------------|--------------------|
| 10 | 532.5 | 16.0 | 855 | 7 | US-11-253-869-2 | Sequence 2, Appl1 |
| 11 | 532.5 | 16.0 | 855 | 7 | US-11-303-608-1 | Sequence 1, Appl1 |
| 12 | 525.5 | 15.8 | 255 | 7 | US-11-254-185-6 | Sequence 6, Appl1 |
| 13 | 525.5 | 15.8 | 255 | 7 | US-11-254-185-38 | Sequence 38, Appl1 |
| 14 | 525.5 | 15.8 | 255 | 7 | US-11-253-869-6 | Sequence 6, Appl1 |
| 15 | 525.5 | 15.8 | 255 | 7 | US-11-253-869-38 | Sequence 38, Appl1 |
| 16 | 522 | 15.7 | 531 | 6 | US-10-522-668-2 | Sequence 2, Appl1 |
| 17 | 522 | 15.7 | 812 | 7 | US-11-318-939-7 | Sequence 7, Appl1 |
| 18 | 520.5 | 15.6 | 423 | 6 | US-10-196-749-320 | Sequence 320, App |
| 19 | 520.5 | 15.6 | 423 | 7 | US-11-101-316-106 | Sequence 106, App |
| 20 | 520.5 | 15.6 | 423 | 7 | US-11-376-673-106 | Sequence 106, App |
| 21 | 515.5 | 15.5 | 311 | 7 | US-11-327-490-41 | Sequence 41, Appl1 |
| 22 | 510.5 | 15.3 | 343 | 7 | US-11-359-554-3 | Sequence 3, Appl1 |
| 23 | 506 | 15.2 | 338 | 7 | US-11-318-939-10 | Sequence 10, Appl1 |
| 24 | 505 | 15.2 | 276 | 7 | US-11-327-490-35 | Sequence 35, Appl1 |
| 25 | 502.5 | 15.1 | 812 | 7 | US-11-318-939-12 | Sequence 12, Appl1 |
| 26 | 498.5 | 15.0 | 339 | 7 | US-11-404-745-1 | Sequence 1, Appl1 |
| 27 | 493.5 | 14.8 | 790 | 7 | US-11-350-703-1 | Sequence 1, Appl1 |
| 28 | 493.5 | 14.8 | 791 | 7 | US-11-318-939-6 | Sequence 6, Appl1 |
| 29 | 493 | 14.8 | 333 | 7 | US-11-318-939-8 | Sequence 8, Appl1 |
| 30 | 492.5 | 14.8 | 810 | 7 | US-11-431-526-2 | Sequence 2, Appl1 |
| 31 | 492.5 | 14.8 | 810 | 7 | US-11-431-663-2 | Sequence 2, Appl1 |
| 32 | 486.5 | 14.6 | 790 | 7 | US-11-318-939-13 | Sequence 13, Appl1 |
| 33 | 485.5 | 14.6 | 311 | 7 | US-11-359-858-2 | Sequence 2, Appl1 |
| 34 | 480.5 | 14.4 | 262 | 7 | US-11-359-554-2 | Sequence 2, Appl1 |
| 35 | 479 | 14.4 | 245 | 7 | US-11-254-185-36 | Sequence 36, Appl1 |
| 36 | 479 | 14.4 | 245 | 7 | US-11-253-869-36 | Sequence 36, Appl1 |
| 37 | 479 | 14.4 | 275 | 6 | US-10-530-798-24 | Sequence 24, Appl1 |
| 38 | 478 | 14.3 | 275 | 6 | US-10-530-798-25 | Sequence 25, Appl1 |
| 39 | 476.5 | 14.3 | 272 | 7 | US-11-327-490-37 | Sequence 37, Appl1 |
| 40 | 475.5 | 14.3 | 272 | 7 | US-11-359-554-6 | Sequence 6, Appl1 |
| 41 | 474.5 | 14.2 | 655 | 7 | US-11-242-617-1 | Sequence 1, Appl1 |
| 42 | 473 | 14.2 | 275 | 6 | US-10-530-798-26 | Sequence 26, Appl1 |
| 43 | 472.5 | 14.1 | 810 | 7 | US-11-318-939-11 | Sequence 11, Appl1 |
| 44 | 470.5 | 14.1 | 272 | 7 | US-11-327-490-40 | Sequence 40, Appl1 |
| 45 | 465 | 14.0 | 275 | 6 | US-10-530-798-23 | Sequence 23, Appl1 |

ALIGNMENTS

RESULT 1
US-10-196-749-64
; Sequence 64, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC340
; CURRENT APPLICATION NUMBER: US/10/196, 749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24


```

Db      55 LeuGlyllellealeuileleuAlaleuAllelleGlyleuGlylleHisPheAspCys 74
QY      560 TGGCAGGGCCCAACAGGATCAGGTACAGAGAGAGAGAGAGCTGTCCCAAGACAGCT 619
Db      75 -----SerGlyLysTyrArgCysArgSerSerPheLysCysIleGluLeuIle 90
QY      620 GTTTCGTGATGACGGGGGTGTGAGACTGCAAGCTGAAGAGTACAGAGCTGGGCTGTAGG 679
Db      91 ThrArgCysAspGlyValSerAspCysLysAspGlyValLysAspGlyValArgCysValArg 110
QY      680 TTTGATGAGGACAAAGCTCTGTTAAATCTACTGGTCTCTCCATCAGTGGCTTCCC 739
Db      111 ValGlyValGlnAsnAlaValleuGlnValPheThrAlaIleHisSer-----TrpLysThr 128
QY      740 ATCTGTAGCAGCACTGGATGACTCTCTACTCAGAGAAGACCTCCGACAGCTGGGCTTCC 799
Db      129 MetCysSerAspAspTrpLysGlyHisTyrAlaAsnValAlaCysAlaGlnLeuGlyPhe 148
QY      800 GAGAGTGTCTACCGGACCGGACCGAGGTTGCC-----CACAGG 835
Db      149 ProSerTyrValSerSerAspAsnLeuArgValSerSerLeuGlnGlyGlnPheArgGlu 168
QY      836 GATTTTGGCAACAGCTTTCATCTTGAATACAC-----TCCACCATCCAGAAAGC 889
Db      169 GluPheValSerIleAspHisIleLeuLeuProAspAspLysValThrAlaLeuHisIleSer 188
QY      890 CTCACAC---AGGTCGTAATGCCCTTCCAGCGGATATTCCTCCGCTGCTTCCCATCTGC 946
Db      189 ValTyrValArgGlnGlyCysAlaSerGlyHisValValThrLeuGlnCysThrAlaCys 208
QY      947 GGA---CTGAGGGCGATACCGGGCGATGCTGGAGGGGGCGCTGGCGATGAGAG 1003
Db      209 GlyHisArgArgGlyTyrSerSerArgIleValGlyLysAsnMetSerLeuLeuSerGln 228
QY      1004 TGGCGCTTGGCAAGTGAAGTGTGACTTGTGAGCAACCCACATCTGTGTGAGGACAGCTCAT 1063
Db      229 TrpProTrpGlnAlaSerIleuGlnPheGlnGlyTyrHisIleuGlySerGlyValIle 248
QY      1064 GACGCCAGTGGGTGCTCATCTGCCGCCACTGCTTCTTCTGTAACCGGAGAAAGTCTCTG 1123
Db      249 ThrProLeuTrpIleIleThrAlaAlaHisCysValTyrAspLeu-----TyrLeuPro 266
QY      1124 GAGGGCGTGAAGGTGTAGCGGGGACAGGACAGGACAGGATGTCGAGGAGCGCTCC 1183
Db      267 LysSerTrpThrIleGlnValGlyLeuValSerLeuLeuAspAsnProAlaProSerHis 286
QY      1184 ATTGCCGAG---ATCATCATCAACAGCAATTAACCCGATGAGAGAGCACTATGATCATC 1240
Db      287 LeuValGluValIleValTyrHisSerLysTyrLysProLysArgLeuGlyAsnAspIle 306
QY      1241 GCCCTCATGCGGCTGTCCAAAGCCCTTACCTGCTCCGCTCATCATCCCTGCTTGCCTC 1300
Db      307 AlaLeuMetLysLeuAlaGlyProLeuThrPheAsnGluMetIleGlnProValCysLeu 326
QY      1301 CCCATGATGAGACAGCATCTTAAGCTCATGAGACCTGATGACAGGCTTGGGAG 1360
Db      327 ProAsnSerGlnLysAsnPheProAspGlyLysValCysTrpTrpSerGlyTyrGlyAla 346
QY      1361 ACCAGGAGACAGATGACAGACATCCCTTCTCCGAGAGTGCAGGTCAATCTCATC 1420
Db      347 ThrGluAspGlyGlyAsp---AlaSerProValLeuAsnHisAlaIleAlaProLeuIle 365
QY      1421 GACTTCAGAAATGCAATGACTACTGTGTTATGACAGTTACTTACCCCAAGATGATG 1480
Db      366 SerAsnLysIleCysAsnHisArgAspValTyrIleGlyIleIleSerProSerMetLeu 385
QY      1481 TGTCTGAGGAGCTTCTGAGGAGAGAGACTCTCCAGAGGAGACAGCGGGGCTCTT 1540
Db      386 CysAlaIleTyrLeuThrGlnGlyValAlaSerSerCysGlnGlyAspSerGlyValProLeu 405
QY      1541 GTCTGTGAGCAAGAACCGCTGTACTCTGACAGGTGTACACAGCTGGGAGCAAGCTGT 1600
Db      406 ValCysGlnGlnLysArgArgLeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys 425

```

```

QY      1601 GGCCAGAAACAAACCTGTGTGTACACCAAGTGCAGAAAGTTCTTCCCTGGATTAC 1660
Db      426 AlaGluValAsnLysProGlyValTyrThrArgValThrPheLeuAspTrpIleHis 445
QY      1661 AGCAAGATGAGAGCGAGCTGCGA 1684
Db      446 GluGlnMetGlnLysArgAspLeuLys 453

RESULT 3
US-10-196-749-330
; Sequence 330, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P9430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 330
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-330

Alignment Scores:
Pred. No.: 4,57e-34 Length: 432
Score: 694.00 Matches: 151
Percent Similarity: 51.7% Conservative: 73
Best Local Similarity: 34.9% Mismatches: 161
Query Match: 20.8% Indels: 48
DB: 6 Gaps: 11

US-10-806-370-11 (1-1748) x US-10-196-749-330 (1-432)
QY      455 CTGCGCAAGTTCACCTGTGCGGAGGCGCAGAGAGCTACCGGTATCGGCTGCTC 514
Db      24 IleProMetGlnLysThrPheArg-----LysValGlyIleProIleIleAlaLeuLeu 41
QY      515 CTCCTCATTCCTCGTGTGTTGCTCATCATCTC-----TTCAG 556

```

```

Db      42 SerLeuAlaSerIleIleValValIleuIleValIleuAspIleValIleuVal 61
QY      557 TTTCGGGAGGCGCCACAGAGGATGATGACAGAGAGGAGGAGACTGTCACCAAC 616
Db      62 PheLeuCySgIyGln-----ProLeuHis 69
QY      617 GCTGTT-----CGCTGTACGGGGTGTGAGCTGCAAGCTGAAGAGTGACGAG 664
Db      70 PheIleProArgIySgIleuCyAspIyGlnIleuAspIyCysProLeuGlnAspGln 89
QY      665 CTGGGGCTGC-----GTGAGCTTTGACTGGGAC 691
Db      90 GluHisCySValIySerPheProGluGlyProAlaValAlaValArgLeuSerIyAsp 109
QY      692 AAGCTCTGCTTAATCTACTGCGGTCCTCCCATGAGGCTTCCCATCTGAGACG 751
Db      110 ArgSerThrIleuGlnIleuAlaAspSerAlaThrIleGlyAsnTrpPheSerAlaCySPhAsp 129
QY      752 AACTGATGATCTCTCTACTCAGAGAACCTGCGAGAGCTGGGTTTGAAGAGTCTCAC 811
Db      130 AsnThrThrGlnAlaLeuAlaGlnThrAlaCySArgIleuMetGlyTrpSer----- 146
QY      812 CGGACAAACCGAGGTTGCCACAGGATTTGCCAACAGCTTCTGAATCTTGAGATACAC 871
Db      147 ArgAlaValGlnIleGlyProAspGlnAspLeuAspValAlaGlnIleThrGluAsnSer 166
QY      872 TCCACCAATCCAGAGAAAGCTCCACAGGTCGAAATGCTCCGAGCGGTATATCTCCTC 931
Db      167 GlnGluIleuArgMetArgAsnSerSerGlyProCySLeuSerGlySerLeuValSerIleu 186
QY      932 CAGTGTTCCTCCACTGCGAGCTGAGGCGCATGACCGGCGGATCTGTGGAGGAGGCGCTGACC 991
Db      187 HisCySLeuAlaCySgIyIySerIleuIySerThrProArgValAlaGlnIyGlnGlnAla 206
QY      992 TCGGATAGCAAGTGGCTTGGCAAGTGAAGTGTGCACTTGGCACACCCATCTGTGGA 1061
Db      207 SerValAspSerTrpProIleGlnValSerIleGlnIyAspIySgIleValIySgIy 226
QY      1052 GGCAGCGCTCATATGACCGCCAGTGGGTGCTACGTGCGGCGCATCTCTTGTGACCGCG 1111
Db      227 GlySerIleuAspProIleTrpValIleuThrAlaAlaHisCySPh-----ArgIyS 244
QY      1112 GAGAGAGTCTCGAGAGGCTGGAAGTGTACCGGCGGACAGCAACCTGACCAAGTTCCT 1171
Db      245 HisThrAspValPheAsnTrpIyValArgAlaGlySerAspIySgIySerPhePro 264
QY      1172 GAGCAGCTTCATTTCCGAGATCATCATC-----AACAGCAATTACCGGATGAG 1222
Db      265 SerLeuAla--ValAlaIySgIleIleIleIleGlnPheAsnPrometIy----- 280
QY      1223 GAGAGCACTATGATACGCGCTCATGCGGCTGTCACGCGCCGATCGCTGCTGCTCAC 1282
Db      281 ProIyAspAsnAspIleAlaIleuMetIyLeuGlnPheProLeuThrPheSerGlyThr 300
QY      1283 ATCCACCTGCTGCTCTCCCATCATGACAGACCTTTAGCTTCAATGAGACCTGCTG 1342
Db      301 ValArgProIleCySLeuProPhePhAspGlnIleuThrProAlaThrProLeuTrp 320
QY      1343 ATCAGAGCTTTGGCAAGACAGAGGAGACAGATACAAAGACATCCCTTCTCCGGAG 1402
Db      321 IleIleGlyTrpGlyPheThrIySgIleAsnGlnIyGlyIySgIyMetSerAspIleLeuIleGln 340
QY      1403 GTGAGGTCATCATCATGACTTCAAGAAATGCAATGACTTACTGCTGATGACAGTTAC 1462
Db      341 AlaSerValGlnValIleAspSerThrArgCySAsnAlaAspAspAlaIyTrpGlnIyGln 360
QY      1463 CTTACCCCAAGATGATGTGTGCTGTGGGACCTTGTGTGGGCGAGACCTCCGACAGGAG 1522
Db      361 ValThrGlnIySgIyMetMetCySAlaGlnIyIleProGlnIyGlyValAspThrCySgIleGly 380
QY      1523 GACGCGGGGGGCTCTTGTGTGTGAGCAGACAAACGCTGTGACTGTGGACGTTGTACC 1582

```

```

Db      381 AspSerGlyIyProLeuMet---TyGlnSerAspGlnTrpHisValIyGlnIleVal 399
QY      1583 AGCTGGGGGACAGCTGTGGCGAGAAACAACTGGTGTGATACCAAGATGACAGAA 1642
Db      400 SerTrpGlyTrpGlyCySgIyGlyProSerThrProIyValIyTrpIySgIySerAla 419
QY      1643 GTTCTTCCTCGATTTTACAGCAAGATGAGACGAGCTG 1681
Db      420 TyIleuAsnTrpIleTyAsnValTrpIySgAlaGlnIleu 432

RESULT 4
US-11-101-316-112
: Sequence 112, Application US/11101316
: Publication No. US20060099657A1
: GENERAL INFORMATION:
: APPLICANT: Goddard, Audrey
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Guirney, Austin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
: TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
: FILE REFERENCE: P323OR1C17C1
: CURRENT APPLICATION NUMBER: US/11/101,316
: PRIOR FILING DATE: 2005-04-06
: PRIOR APPLICATION NUMBER: 10/063526
: PRIOR FILING DATE: 2002-05-03
: PRIOR APPLICATION NUMBER: 10/006867
: PRIOR FILING DATE: 2001-12-06
: PRIOR APPLICATION NUMBER: PCT/US00/23328
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: 09/380137
: PRIOR FILING DATE: 1999-08-25
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: NUMBER OF SEQ ID NOS: 170
: SEQ ID NO 112
: LENGTH: 432
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-101-316-112

Alignment Scores:
Pred. No.: 4,57e-34 Length: 432
Score: 694.00 Matches: 151
Percent Similarity: 51.7% Conservative: 73
Best Local Similarity: 34.9% Mismatches: 161
Query Match: 20.8% Indels: 48
DB: Gaps: 11

US-10-806-370-11 (1-1748) x US-11-101-316-112 (1-432)
QY      455 CTGCGCAAGTTTACCTTGGCGGAGGCGCAGAGCACTGACCGTCAATCGGAGTGCCTC 514
Db      24 IleProMetGlnThrPheArg-----IySgValIyGlyLeProIleIleIleAlaLeuIleu 41
QY      515 CTCTCATTTGCCCTCGTGGTTCCTCATCATCTC-----TTCAG 556
Db      42 SerLeuAlaSerIleIleValValIleuIleValIleuAspIySerIyTrp 61
QY      557 TTTCGGGAGGCGCCACAGAGGATGATGACAGAGAGGAGAGCTGTCCCAACGAC 616
Db      62 PheLeuCySgIyGln-----ProLeuHis 69
QY      617 GCTGTT-----CGCTGTACGGGGTGTGAGCTGCAAGCTGAAGAGTGACGAG 664
Db      70 PheIleProArgIySgIleuCyAspGlyGlnIleuAspCysProLeuGlnIyGlnAspGln 89
QY      665 CTGGGGCTGC-----GTGAGCTTTGACTGGGAC 691
Db      90 GluHisCySValIySerPheProGluGlyProAlaValAlaValArgLeuSerIyAsp 109

```



```
Db 346 AspAlaCysGlnGlyAspSerGlyGlyProPheValCysGlnAspSerIleSerGlyThr 365
QY 1556 AACCGCTGTAAGCTGGCAGGTGTACACAGCTGGGCGACAGGCTGTGGCCGAGAAACAA 1615
Db 366 SerArgTrpArgLeuGlyGlyIleValSerTrpGlyThrGlyCysAlaLeuValArgGly 385
QY 1616 CCTGGGTGTACACCAAGTACAGAAAGTCTTCCCGATTTACAGCAGATGAGAGC 1675
Db 386 ProGlyValIleThrIleValIleAspPheArgGlnTrpIlePheValIleValIleValIle 405

RESULT 7
US-11-293-697-3002
; Sequence 3002, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3002
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3002

Alignment Scores:
Pred. No.: 4,29e-25 Length: 305
Score: 543.50 Matches: 121
Percent Similarity: 54.4% Conservative: 39
Best Local Similarity: 41.2% Mismatches: 107
Query Match: 16.3% Indels: 28
DB: Gaps: 9

US-10-806-370-11 (1-1748) x US-11-293-697-3002 (1-305)
QY 873 CCACCATCCAGAAAGCTCCACAGGTGTGATGCCCTCCAGCGGTATATCTCCCTCC 932
Db 8 ProProArgAlaGlyProSerLeuGlySerGlyAspLeu-----GlySerProLeu 25
QY 933 AGTGTCCCA-----TGCGACTGAGGCGCATG---ACCGG 967
Db 26 SerProProProAlaAspProCysProThrAspCysGlyArgArgIleuProValAsp 45
QY 968 CGGATGCTGGAGGCGGCTGGCGCTCGGATACAGAGTGGCTTGGCAAGTGTGTCAC 1027
Db 46 ArgIleValGlyGlyArgAspThrSerLeuGlyArgTrpProTrpGlnValSerLeuArg 65
QY 1028 TTGGGACACCAACATCTGTGAGGACGCTGATTAAGCCACAGTGGGTCTACATGCC 1087
Db 66 TyAspGlyAlaIleuGlyGlySerLeuSerGlyAspTrpValIleuThAla 85
QY 1088 GCCCATGCTTTCTTGTGACCGGAGAAAGTCTGGAGGCTGGAAGTGTACGCGGG 1147
Db 86 AlaHieGly---PheProGlnArgGlnArgValIleuSerArgTrpArgValPheAlaGly 104
QY 1148 ACC-----AGCAACTGACACCACTGCTGAGGACGCTTCATGCGGAGATATC 1198
Db 105 AlaValAlaGlnIleAspProHISGlyLeu-----GlnIleuGlyValGlnAlaValAl 122
QY 1199 ATCAACAGCAATTAC-----ACCGATGAGGAGGAGCACTATGACATC 1240
Db 123 TyHISGlyGlyTyLeuProPheArgAspProAsnSerGlnGlnuAsnSerAlaAspIle 142
QY 1241 GCCCTTATGCGGCTGTCCAAAGCCCTGACCTGTCCGCTACATCAACCTGCTGCTC 1300
Db 143 AlaLeuValHISerSerSerProLeuProIleuThrGlnIleGlnProValCysLeu 162
```

```
QY 1301 CCATCATGACAGACAGCTTTAGCCTCAATGAGACCTGTGATACAGGCTTTGGCAG 1360
Db 163 ProAlaIleGlyGlnIleuValaAspGlyGlyIleCysThrValIleGlyTrpGlyAsn 182
QY 1361 ACCAGGAGACAGATGACAAAGATCCCTTCTCCGAGAGTGAAGTCAATTCATC 1420
Db 183 ThrGlnTyTrpGlyGlnGlnAlaGly---ValLeuGlnGlnIleuArgValProIleIle 201
QY 1421 GACTTCAAGAAATGCAATGACTTGTGCTATGACAGTACTTACCACCAAGATGATG 1480
Db 202 SerAsnAspValCysAsnGlyAlaAspPheTyGlyAsnGlnIleuSerProIleuPhe 221
QY 1481 TGTGCTGGGACCTGTGAGGAGCAGACACTCTGCGAGGAGACAGCGGAGGCTCTT 1540
Db 222 CysAlaGlyTyTrpGlnGlyGlyIleAspAlaCysGlnIleuAspSerGlyGlyProPhe 241
QY 1541 GTCTGTGAG-----CAGAACACCGCTGTACTGCGAGGTGTACACAGCTGG 1588
Db 242 ValCysGlnAspSerIleSerArgTrpProArgTrpArgLeuGlyIleValSerTrp 261
QY 1589 GGCACAGCTGTGCGCAGAGAAACAAACCTGTGTGTACCAAGACAGAAAGTCTT 1648
Db 262 GlyThrGlyCysAlaLeuAlaGlnIleuSerProGlyValIleThrIleValSerAspPheArg 281
QY 1649 CCTGATTTACAGCAAGATGAGAGCGAGGTGCGATTGCA 1690
Db 282 GlnTrpIlePheGlnAlaIleValIleValIleValIleValIleValIleValIle 295

RESULT 8
US-10-530-187-242
; Sequence 242, Application US/10530187
; Publication No. US20060183120A1
; GENERAL INFORMATION:
; APPLICANT: THE, Bin Team
; APPLICANT: TAKAHASHI, Masayuki
; TITLE OF INVENTION: Molecular Subclassification of Kidney Tumors and the Discovery of
; FILE REFERENCE: 28927-0014
; CURRENT APPLICATION NUMBER: US/10/530,187
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US2003/031476
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: US 60/415,775
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 242
; LENGTH: 853
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-530-187-242

Alignment Scores:
Pred. No.: 2,03e-24 Length: 853
Score: 532.50 Matches: 163
Percent Similarity: 40.0% Conservative: 61
Best Local Similarity: 29.1% Mismatches: 175
Query Match: 16.0% Indels: 161
DB: Gaps: 25

US-10-806-370-11 (1-1748) x US-10-530-187-242 (1-853)
QY 282 TTTCAGGTCCTCATCCGCGAGTCAATCATCCGCGAGTCAAGCTCGTGACAACTCC 341
Db 349 PheAsnSerProTyTrpProGlyHISGlyTy-----ProPro 360
QY 342 CAACAGAGTGTACTTTAGAGCAACACCACTGGGAGGCTGTACCATTCATCTTC 401
Db 361 AsnIleAspCysThr-----TrpAsnIleGlnValIleProAsnAsn--- 373
QY 402 CTGCGAGTCAAGAC-----CAGCAACAGAGGCA 431
Db 374 -----GlnHISValIleValArgPheValPhePheTyLeuLeuGlnProGlyVal 390
```

```

QY 432 CCAAGGAGAGCCGAGTACGAGCCTGCCAAGTTCA----- 467
   |||
   |||
Db 391 Pro-----AlaGlyThrCysProLysAspTyrValGluIleAsnGlyGluLys 406
QY 468 -----CCTGGCCGGAGAGGCCAGA----- 485
   |||
   |||
Db 407 TyrCysGluIleArgSerGlnPheValValThrSerAsnSerAsnLysIleThrValArg 426
QY 486 -----ACAGCTACCGCTCATCGGGGTGCGTCTCTCTCATTTGCTGGTGG 533
   |||
   |||
Db 427 PheHisSerArgLysIleThrAspThrGlyPheLeuAlaGluTyrLeu----- 443
QY 534 TTTCGCTCATCATCTCTTTCAGATTTCGAGGCCACACAGAGATGACGTAAGAAGAC 593
   |||
   |||
Db 444 -----SerTyrAspSerSerAspProCysProGluGlnPheThrCysArg 458
QY 594 AGAGGAGAGAGTCTCCAGACGAGCTGTTCGTGTGACGGGGGTGTGACTGCAGCTGA 653
   |||
   |||
Db 459 ThrGlyArg--CysIleArgGluLeuArgCysAspGlyTyrPalaAspCysThrAspH 478
QY 654 AGAGTGAAGAGTGGGTGGCTGAGGTTTACCTGGACAAGCTCTGCTTAAATCTACT 713
   |||
   |||
Db 478 IAspSerArgLysLeuAsnCys-----Serc 486
QY 714 CTGGGCTCTCCCATCAG-----TGCGTTC 737
   |||
   |||
Db 486 LysAspAlaGluIleGlnPheThrCysLysAsnLysPheCysLysPheProLeuPheThr 504
QY 738 CCATCTGTAGAGC-----AACTGGAATGACTCTTACTACAGAAAGACCTGC----- 784
   |||
   |||
Db 505 --ValCysAspSerValAsnAspCysGluAspAsnSerAspGluGlnGlyCysSerCysP 524
QY 785 -----CAGACGCTGG 794
   |||
   |||
Db 524 LysIleGlnIlePheArgCysSerAsnGlyLysCysLeuSerLysSerGlnGlnCysAsnG 544
QY 795 GTTTCGAGAGTGTCAACCGGACAAACGAGGTGGCCACAGGAAATTTGGCAACGCTTCT 854
   |||
   |||
Db 544 LysLysAspAspCysGluAspGlySerAspGluIleAspCysProLysValAlaAsnValAl 564
QY 855 CAATCTTGAGATCAACTCCACCAATCCAGAAAGCTCCACAGGCT-----G 902
   |||
   |||
Db 564 hTrCysThrLysHisIleThrTyrArgCysLeuAsnGlyLeuCysLeuSerLysGluAsnProG 584
QY 903 AATGCCCTTCCACGCGATATCTCCCTCCAGTGTCC----- 940
   |||
   |||
Db 584 LysCysAspGlyLysGlu-----AspCysSerAspGlySerAspGlyLysAspC 600
QY 941 --CACTGCGAGTGAAGGCCATGAC-----GGGCGGATGTGAGAGGGCGCTGGCT 992
   |||
   |||
Db 600 LysAspCysGluLeuArgSerPheThrArgGlnAlaArgValAlaGlyLysThrAspAla 620
QY 993 CGGATGACAGTGGCTTTGGCAAGTGAAGTGTGAC--TTGGGACCAACCCACATCTGTG 1049
   |||
   |||
Db 620 spGluGlnGluTyrProTyrGlnValSerLeuHisAlaLeuGlnGlnGlyHisIleCysG 640
QY 1050 GAGGCAAGCTATTAAGGCGCCAGTGGGTGCTCACTGGCGGCCCATGCTCTCTGTAACC 1109
   |||
   |||
Db 640 LysAlaSerLeuIleSerProAsnIleValSerAlaAlaHisCysTyrIleAspAsp 660
QY 1110 GCGAGAGAGTCTCGAGAGGC-----TGGAAGGTGTACGCGG 1145
   |||
   |||
Db 660 rg-----GlyPheArgTyrSerAspProThrGlnIleThrPheIleAsnLeuG 675
QY 1146 GCACCAAGACCTGACACGATTGCTTGAAGCACCTCC-----A 1184
   |||
   |||
Db 675 Lys-----LeuHisArgLysSerGlnArgSerAlaProGlyValGlnGlnLysArgL 692
QY 1185 TTGCGGAGATGATGATTAACAGCAATTACCGAGTAGAGAGAGAGATGATGACATGGCC 1244
   |||
   |||
Db 692 euLysArgIleIleSerHisPheProPheAsnAspPheThrPheAspTyrAspIleAla 712

```

```

QY 1245 TCATGCGGCTGTCCAAAGCCCTTGAACCTGTCCGCTCAATCCACCTGTGCTCCCA 1304
   |||
   |||
Db 712 euLeuGluLeuGluLysProAlaGluTyrSerSerMetValArgProIleCysLeuProA 732
QY 1305 TGCAATGACAGACCTTTAGAGCTCAATGAGACCTGTGATTCACAGGCTTGGCAAGCA 1364
   |||
   |||
Db 732 spAlaSerHisValPheProAlaGlyLysAlaIleTyrValHisIleThr 751
QY 1365 GGGAGACAGATGACAAAGACATCCCTTCTCCGGGAGGTGACAGTCAATCTCACT 1424
   |||
   |||
Db 752 --GlnTyrGlyGlyThrGlyAlaLeuIleLeuGlnLysGluIleAlaArgValIleAsnG 771
QY 1425 TCAAGAAATGCAATGACTACTTGTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTG 1484
   |||
   |||
Db 771 LysThrCysGluAsnLeu-----ProGlnGlnIleThrProArgMetMetCysAv 789
QY 1485 CTGGGAGACCTTCTGTGGGGGCGAGAGACTCTGCCAGGAGACAGGGGGGCTCTT--G 1541
   |||
   |||
Db 789 AlGlyPheLeuSerGlyValAlaAspSerCysGlnGlyAspSerGlyLysProLeuSer 809
QY 1542 TCTGTGACAGAACAAACCGCTGTGACTCTGCAGAGTGTCAACAGCTGGGCGACAGCTGTG 1601
   |||
   |||
Db 809 euValGluAlaAspGlyArgIlePheGlnAlaGlyValAlaSerTyrGlyAspGlyCysA 829
QY 1602 GCCAGAGAAACAAACCTGTGTGTGTACACCAAGTGCAGAGAAATTTCTTCCTGAT 1657
   |||
   |||
Db 829 LysGlnArgAsnLysProGlyValIleTyrThrArgLeuProLeuPheArgAspTyrIle 847

```

RESULT 9

US-11-254-185-2

Sequence 2, Application US/11254185

Publication No. US2006009625A1

GENERAL INFORMATION:

APPLICANT: CRAIK, CHARLES S.

APPLICANT: TAKEUCHI, TOSHIIKO

APPLICANT: SCHUMAN, MARC

TITLE OF INVENTION: WT-SPI SERINE PROTEASE

FILE REFERENCE: 28644-701.302

CURRENT APPLICATION NUMBER: US/11/254,185

CURRENT FILING DATE: 2005-10-18

NUMBER OF SEQ ID NOS: 83

SOFTWARE: PatentIn version 3.3

SEQ ID NO 2

LENGTH: 855

TYPE: PRT

ORGANISM: Homo sapiens

US-11-254-185-2

Alignment Scores:

| Pred. No.: | Length: | Matches: |
|------------------------|---------|------------------|
| 532.50 | 855 | 163 |
| Percent Similarity: | 40.0% | Conservative: 61 |
| Best Local Similarity: | 29.1% | Mismatches: 175 |
| Query Match: | 16.0% | Indels: 161 |
| | Gaps: | 25 |

US-10-806-370-11 (1-1748) x US-11-254-185-2 (1-855)

```

QY 282 TTTCAGAGTCTTATCCGCGAGTCAATCCGCGAGTCAAGCTCGGTGACAACTCCC 341
   |||
   |||
Db 351 PheAsnSerProTyrTyrProGlyHisIleTyr-----ProPro 362
QY 342 CAACCAAGTGTACTTTAGAGCAACACAGTGGGGGTGTACCATCGATCATCTC 401
   |||
   |||
Db 363 AsnIleAspCysThr-----TrpAsnIleGluValProAsnAsn--- 375
QY 402 CTGCGAGTACGAC-----CAGCAACCGGGCCA 431
   |||
   |||
Db 376 -----GlnHisValLysValArgPheLysPhePheTyrLeuLeuGluProGlyVal 392
QY 432 CCAAGGAGAGCCAGGTACGAGCTGCGCAAGTTCA----- 467
   |||
   |||
Db 393 Pro-----AlaGlyThrCysProLysAspTyrValGluIleAsnGlyGluLys 408

```

| | | | |
|----|------|--|------|
| QY | 468 | -----CCTGACGGGAGGCGCAG----- | 485 |
| Db | 409 | TyrCysGlyGlnAArgSerGlnPheValValThrSerAsnSerAsnIleThrValArg | 428 |
| QY | 486 | -----AGACGTACCGGTCATCGGGTGGCTTCCTCTCAATTGSCCTGGTG | 533 |
| Db | 429 | PheHisSerAspGlnSerTyrThrAspArgIlePheLeuAlaGluTyrLeu----- | 445 |
| QY | 534 | TTTGGCTCATCATCTCTTCACAGTTTTCGGACGGGCCACACAGGATCATGATCAAGAGGC | 593 |
| Db | 446 | -----SerTyrAspSerSerAspProCysProGluGlnPheThrCysArg | 460 |
| QY | 594 | AGAGGAGAGAGCTGTCCCAAGCAGCCTGTTCAGCTGACGGGGGTGGATGACGTCAAGCTGA | 653 |
| Db | 461 | ThrIleArgArg-CysIleArgGlySerLeuLeuAlaGlyCysAspGlyTTPAlaAspCysThrAspH | 480 |
| QY | 654 | AGAGTGAACGAGCTGGGCTGGCTGAGGTTTGACTGGGACAACTCTTCGTTAAATTAATTA | 713 |
| Db | 480 | IsSerAspGlnLeuAlaAsnCys----- | 488 |
| QY | 714 | CTGGGCTCTCCATCAG----- | 737 |
| Db | 488 | yAspAlaGlnAArgIlePheThrCysValAsnIlyPheCysValProLeuPheTrp----- | 506 |
| QY | 738 | CCATCTTGAACAGC-----AACTGGAATGACTCTTACTAGAGAAGACTGC----- | 784 |
| Db | 507 | --ValCysAspSerValAlaAspCysGlyAspAsnSerAspGlnGlnIlyCysSerCysP | 526 |
| QY | 785 | -----CAGCAGCTGG | 794 |
| Db | 526 | roAlaGlnThrPheArgCysSerAsnGlnIlyCysLeuSerIysSerGlnIlnCysAsnG | 546 |
| QY | 795 | GTTTCGAGAGTGCACCCGACACACCGAGGTGGCCACAGGGAATTTTCCCAACAGCTTCT | 854 |
| Db | 546 | IlyAspAspAspCysGlyAspIlyAspIlyAspArgIlyAlaSerCysProIlyAlaAsnValValT | 566 |
| QY | 855 | CAATCTTGAGATCAACTCCACCATCCACGAAAGCTTCACAGGCT-----G | 902 |
| Db | 566 | hrCysThrIlyAsnIleThrTyrArgCysLeuAsnGlnIlyLeuCysLeuSerIysGlyAsnProG | 586 |
| QY | 903 | AATGSCCTTCCCAACGGATATCTCCCTCCAGTGTCC----- | 940 |
| Db | 586 | IlyCysAspArgIlyValSerIln-----AspCysSerAspIlySerAspGlnIlyAspC | 602 |
| QY | 941 | --CACTCGGAGTGAAGGCGCATGAC-----GGGCGGATGTGGAGAGGCGCTGCT | 992 |
| Db | 602 | yAspCysGlyIleuArgSerPheThrArgGlnAlaArgValAlaGlyIlyThrAspAlaA | 622 |
| QY | 993 | CGGATGACAAAGTGGCTTTGGCAAGTAGTCTGCAC--TTGGGACACACCATCTGTG | 1049 |
| Db | 622 | spGlnGlyGlnTrpProTyrPheIlnAlaSerLeuAlaLeuGlnIlnIlyHisIleCysG | 642 |
| QY | 1050 | GAGGCGAGCTCATTTGACGCGCAGGAGGGGTGCATCGGCGCCACATGCTTTCGTAACC | 1109 |
| Db | 642 | IlyAlaSerLeuIleSerProAsnTrpLeuValSerAlaAlaHisCysTyrIleAspAspA | 662 |
| QY | 1110 | GGGAGAAAGTCTGAGAGGC-----TGGAAAGTATACGGG | 1145 |
| Db | 662 | rg-----GlyPheArgTyrIserAspProThrGlnTrpThrAlaPheLeuG | 677 |
| QY | 1146 | GCACCGACAACTTGACACAGTTGGCTGAGGACGCTCC-----A | 1188 |
| Db | 677 | Ily-----LeuHisAspGlnSerGlnArgSerAlaProGlyValGlnGlnArgArgI | 694 |
| QY | 1185 | TTGGCGAGATCATCATCAACAAATTACACCGATGAGAGAGACGACTATGACATCGGCC | 1244 |
| Db | 694 | eulysArgIleIleSerHisProPhePheAsnAspPheThrPheAspIlyAspIleAlaI | 714 |
| QY | 1245 | TCAATCGGGCTGTCCAAAGCCCTGAACCCGTGTCGGCATCAACCCGTGCTTGGCTCCCA | 1304 |
| Db | 714 | eulSeuGlnLeuGlnIlySerProAlaGlnIlyTyrSerSerMetValArgProIleCysLeuProA | 734 |

```

OY      1305  TGCATGACACACCTTTTAACTTCATATGAGACCTCTGGATTCACAGGCTTTGGACAAGCA 1366
Db      734  sp1aSerHisValAlpheProAlaelyVala1a1eTPrValThelYTrpPolynHsthr- 753
OY      1365  GGGAGACAGATGACACAGACATCCCTCTCTCCGGAGGTGCAGGTCAATCTGACT 1424
Db      754  --GlnTYrGlyGlyThrGlyAlaLeuileuLeuGlnlyGlyGluileargValilleaeng 773
OY      1425  TCAAGAAATGCATAGACTACTTGCTGTATGACAGTACTTACCTTACCCCAAGATGATGTGTG 1488
Db      773  InThrThrCysGluAsnLeuLeu-----ProGlnGlnIleThrProArgMetMetCysV 791
OY      1485  CTGGGAGACCTTCGGGGGGGAGAGACTCTCTCCAGGGAGACAGGGGGGCGCTCTT---G 1544
Db      791  a1GlyPheLeuSerGlyGlyAlaAspSerCysGlnGlyAspSerGlyGlyProLeuSers 811
OY      1542  TCTGTGACAGCAACAACCGCTGTACTGTGACGTGCAGGCTGTACACAGCTGGGGCACAGCTGTG 1604
Db      811  erVala1GluAlaAspGlyArgIlePheGlnAlaGlyAlaVala1SerTrrPolyspGlyCySa 831
OY      1602  GCCAGAGAAACAACCTGTGTGTACACCAACAAAGTGACAGAACTTCTTCCTGGATT 1657
Db      831  laGlnArgAsnlyGlyProGlyValYrThrArgLeuProLeuPheArgAspTrrPile 849

RESULT 10
US-11-253-869-2
; Sequence 2, Application US/11253869
; Publication No. US20060104979A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; APPLICANT: SCHUMAN, MARC
; TITLE OF INVENTION: WT-SP1 POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 28644-701.303
; CURRENT APPLICATION NUMBER: US/11/253.869
; CURRENT FILING DATE: 2005-10-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-253-869-2

Alignment Scores:
Pred. No.: 2.03e-24 Length: 855
Score: 532.50 Matches: 163
Percent Similarity: 40.0% Conservative: 61
Best Local Similarity: 29.1% Mismatches: 175
Query Match: 16.0% Indels: 161
DB: 7 Gaps: 25

US-10-806-370-11 (1-1748) x US-11-253-869-2 (1-855)
OY      282  TTTCAGGTCTTCATCCGGCAGGTCAATCCGCCAGGTGACCTCGGTGCAAACTTCC 341
Db      351  PheAsnSerProTYrTYrProGlyHisIstYr-----ProPro 362
OY      342  CAACCAGAGTACTACTTGTTAGACACACACAGTGGGGGCTGTACCATCCGATCATATC 401
Db      363  AsnIleLeuCysTTr-----TrrPnIleGluValProAsn-- 375
OY      402  CTGCAGGTGCAGCAC-----CAGCAACCAAGGGCCA 431
Db      376  -----GlnHisValIyVala1ArgPheIyPheTherYrLeuLeuGlnProGlyAla 392
OY      432  CCAGGAGAGCCCAAGTACGAGCTCTGCCAAATTCA----- 467
Db      393  Pro-----AlaGlyThrCysProIyAspTyrVala1GluIleAsnGlyGlyIyS 408
OY      468  ---CCTGGCGGAGAGGGCCAG-- 485
Db      409  TyrCysGlyGluArgSerGlnPheVala1ThrSerAsnSerAsnlyS1leThrVala1Arg 428

```

```

OY 486 -----ACAGTACCGCTCATCGGGTGGTGGTCTCTCTCATTTGCGCTGG 533
Db 429 PheHisSerArgInsThrThrAspThrGlyPheLeuAlaGlyTyrLeu----- 445
OY 534 TTTCGCTCATCATCTCTTCCAGTTCTGGAGGGCCACAGGATCAGGTACAGAGC 593
Db 446 -----SerTyrAspSerSerAspProCysProGlyGlnPheThrCysArg 460
OY 594 AGAGGAGAGCTGTCCACAGACGCTGTCTGGACCGGGGTGTGTGATCGCAAGCTGA 653
Db 461 ThrGlyArg-CysAlaGlyGlyLeuArgCysAspGlyTyrAlaAspCysThrAspH 480
OY 654 AGAGTACAGAGCTGGCTGCGTGAAGTTTGAAGTGGACAAAGTCTGTGTTAAATCTACT 713
Db 480 IsSerAspGlnLeuAsnCys-----Serc 488
OY 714 CTGGGCTCTCCATCAG-----TGCTTC 737
Db 488 ysaAraAlaGlnHISGlnPheThrCysLeuAsnLysPheCysLeuProLeuPheTrp--- 506
OY 738 CCATCTGTAGAGC-----AAGTGAATGATCTCTACTGAGAAAGCTGC----- 784
Db 507 --ValCysAspSerValAsnAspCysGlyAspAsnSerAspGlnGlnGlyCysSerCysP 526
OY 785 -----CAGCAGCTGG 794
Db 526 rolaGlnThrPheArgCysSerAsnGlyLysCysLeuSerLysSerGlnGlnCysAsnG 546
OY 795 GTTTCGAGAGTGTCTACCCGACAAACGAGTTGCCACAGGATTTTCCACAGCTTCT 854
Db 546 LysAspAspCysGlyAspGlySerAspGlnAlaSerCysProLysValAsnValAla 566
OY 855 CAATCTGAGATCAATCCACCATCAGGAAAGCTTCCACAGCTCT-----G 902
Db 566 hrCysThrLysHisThrTyrArgCysLeuAsnGlyLeuCysLeuSerLysGlyAsnProG 586
OY 903 AATGCCCTCCAGCGGTATATCTCCCTCCAGCTTCC----- 940
Db 586 LysAspGlyLysGln-----AspCysSerAspGlySerAspGlnLysAspC 602
OY 941 --CACTGCGGATGAGGCGCATGAC-----GGGCGGATGCGGAGGCGGCTGCT 992
Db 602 ysaAspCysGlyLeuArgSerPheThrArgAlaArgValAlaGlyGlyThrAspAla 622
OY 993 CGGATACCAAGTGGCTTTGGAAATGATCTGCAC--TTGCGACACACCATCTGTG 1049
Db 622 spGlnGlyGlnTrpProTrpGlnValSerLeuHisAlaLeuGlyGlnGlyHisLecCysG 642
OY 1050 GAGGACGCTCATTTGACGCGCAGGAGGTGCTACGCGCGCCACGCTTCTTGGTACC 1109
Db 642 LysAlaSerLeuHisSerProAsnTrpLeuValSerAlaHisCysArgLysLeuAspA 662
OY 1110 GCGAGAAAGTCTTGAGAGGC-----TGAGAGGTGTACGGG 1145
Db 662 rg-----GlyPheArgTyrSerAspProThrGlnTrpThrAlaPheLeuG 677
OY 1146 GCACACAGACCTGCACCAAGTTGCTGAGGAGCTCC-----A 1184
Db 677 Lys-----LeuHisAspGlnSerGlnArgSerAlaProGlyValGlnGlnArgArgL 694
OY 1185 TTGCGGAGATCATGATCAACAGCAATTACACCAATGAGAGAGAGCATATGACATGCGCC 1244
Db 694 euLysArgGlnLeuHisSerHisProPhePheAsnAspPheThrPheAspTyrAspLysAla 714
OY 1245 TCATGCGGCTGTCCAAAGCCCTGAACCTGTCCGCTCATCATCAGCTTGTGCTGCCA 1304
Db 714 euLeuGlnLeuGlnLysProAlaGlnLysArgSerMetValArgProLysLeuProA 734
OY 1305 TGCTTGACAGACCTTTAGCTCATAGACAGCTGTGATACAGGCTTTGGCAGAGCA 1364
Db 734 spAlaSerHisValPheProAlaGlyLysAlaLeuTrpValThrGlyTyrPylHisThr- 753

```

```

OY 1365 GCGAGACATGACAAACATCCCTTCCGCGAGGTGACAGTCAATCTCACTGACT 1424
Db 754 --GlnTyrGlyGlyThrGlyValAlaLeuLeuGlnLysGlyGlnLysValAlaLeu 773
OY 1425 TCAAGAAATGCATGACTACTTGTGTTATGACACTTACCTTACCCCAAGATATGTGTG 1484
Db 773 LysThrCysGlnAsnLeu-----ProGlnGlnHisThrProArgMetMetCysV 791
OY 1485 CTGGGAGCTTCCGCGGCGGAGAGACTCTGCGGAGGAGACAGCGGGGGCTCTT---G 1541
Db 791 AlGlyPheLeuSerGlyValAlaAspSerCysGlnGlyAspSerGlyGlyProLeuSer 811
OY 1542 TCTGTGACAGAAACAAACCGCTGTGACTCTGCGAGGTGTACACAGCTGGGCGCAGCTGTG 1601
Db 811 euValGlnAlaAspGlyArgGlyPheGlnAlaGlyValAlaSerTrpGlyAspGlyCysA 831
OY 1602 GCCAGAGAAACAAACCTGTGTGTACACCAAGTGAAGAGAGTTCTTCCCTGGATT 1657
Db 831 LysGlnArgAsnLysProGlyValAlaTyrThrArgLeuProLeuPheArgAspTrpLle 849

RESULT 11
US-11-303-608-1
Sequence 1, Application US/11303608
Publication No. US20060171884A1
GENERAL INFORMATION:
APPLICANT: Folz, Ian
APPLICANT: King, Chadwick
APPLICANT: Ling, Peter Koon Bong
APPLICANT: Kang, Jaepal Singh
APPLICANT: Manchulenko, Kathy
APPLICANT: Chen, Francine
APPLICANT: Scatena, Caroline Darne
APPLICANT: Key, Bruce A.
APPLICANT: Madison, Edwin R.
APPLICANT: Godfrey, Wayne R.
APPLICANT: Morikowski, Stanislaw K.
APPLICANT: Richardson, Jennifer H.
TITLE OF INVENTION: BINDING PROTEINS SPECIFIC FOR HUMAN
TITLE OF INVENTION: MATRIXPHASE
FILE REFERENCE: ABGENIX.148A
CURRENT APPLICATION NUMBER: US/11/303,608
CURRENT FILING DATE: 2005-12-16
PRIOR APPLICATION NUMBER: 60/706467
PRIOR FILING DATE: 2005-08-08
PRIOR APPLICATION NUMBER: 60/637859
PRIOR FILING DATE: 2004-12-20
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 855
TYPE: PR
ORGANISM: Homo sapiens
US-11-303-608-1

Alignment Scores:
Pred. No.: 2,036-24 Length: 855
Score: 532.50 Matches: 163
Percent Similarity: 40.0% Conservative: 61
Best Local Similarity: 29.1% Mismatches: 175
Query Match: 16.0% Indels: 161
DB: 7 Gaps: 25

US-10-806-370-11 (1-1748) x US-11-303-608-1 (1-855)
OY 282 TTTCAGGTCTCTATCCGAGGATCATATCCGACAGTACAGCTCGGTGACAACTGCC 341
Db 351 PheAsnSerProCysTyrTyrProGlnHisTyr-----ProPro 362
OY 342 CAACCAAGTGTACTCTTGTATGAGCAACACAGTGGGGGTGTACCGATCGATCTC 401
Db 363 AsnLysAspCysThr-----TrpAsnLysGlnValProLeuAsn--- 375
OY 402 CTGCCAGTCAAGCAC-----CAGCAACGAGGCGCA 431

```

```

Db      376 -----GlnHisValIysValArgPheIysPhePheThrIleuLeuGluProGlyVal 392
QY      432 CCAGGAGAGCCCGAGTACGAGCTGCCCAATTCA----- 467
Db      393 Pro-----AlaGlyThrCysProIysAspIyrValGluIleAsnGlyIuLys 408
QY      468 -----CCTGGCGGAGGCCAG----- 485
Db      409 TyrCysGlyIuLysSerGlnPheValValThrSerAsnSerAsnIleThrValArg 428
QY      486 -----ACAGCTACCGCTCATCGGGTGCCTGCTCCCTGATTCCTCGGTGG 533
Db      429 PheHisSerAspGlnSerIyrThrAspThrIlePheuAlaGluIyrIleu----- 445
QY      534 TTTCGTCATCATCTCTTCCAGTTCTGGAGGGCCACAGAGGATGACGATGACAGAGC 593
Db      446 -----SerIyrAspSerSerAspProCysProGlyGlnPheThrCysArg 460
QY      594 AGAGGAGAGCTGTCCCAAGCAGCTGTGCTGTGACGGGTGTGAGTGTGCAAGCTGA 653
Db      461 ThrIleIyrG--CysIleArgIuLysLeuArgCysAspGlyTTPAlaAspCysThrAspH 480
QY      654 AGAGTGCAGAGCTGGGCTGAGGTTTGTACGTGGACAAGTCTGTGTTAAATCTACT 713
Db      480 IserAspGlyLeuAsnCys-----Serc 488
QY      714 CTGGGCTCTCCATCAG-----TGCGTTC 737
Db      488 ysaAspAlaGlyHisGlnPheThrCysIysAsnIysPheCysIysPheProIeuPheTrp---- 506
QY      738 CCATGTGTAGCAGC-----AAGTGAATGACTCTTACTAGAGAAGACCTGC----- 784
Db      507 --ValCysAspSerValAsnAspCysGlyAspAsnSerAspGluGlnGlyCysSerCysP 526
QY      785 -----CAGAGCTGG 794
Db      526 roAlaGlnIthrPheArgCysSerAsnGlyIysCysLeuSerIysSerGlnIuLysAsnG 546
QY      795 GTTTCGAGAGTGTCAACCGGACAAACGAGGTGGCCACAGGATTTTCCCAACAGCTTCT 854
Db      546 IlyLysAspAspCysGlyAspGlySerAspGluAlaSerCysProIysValAsnValIalt 566
QY      855 CAATCTTGAGATACACTCCACCATCCAGGAAGACCTTCACAGCTCT-----G 902
Db      566 hrCysThrIysHisThrIyrArgCysLeuAsnGlyIleuCysLeuSerIysIysAsnProG 586
QY      903 AATGCCCTTCCCGAGGATATCTCCCTCCAGTGTCC----- 940
Db      586 IuCyAspAspGlyIysGln-----AspCysSerAspGlySerAspGlyIuLysAspC 602
QY      941 --CACTGCGAGTACAGGCGCATGAC-----GGCGGATGCTGGAGAGGCGCTGCT 992
Db      602 ysaAspCysGlyIleuArgSerPheThrArgGlnAlaArgValIleGlyIthrAspAla 622
QY      993 CGGATGACAGTGGCTTTGGCAAGTGAAGTGTGCAC--TTGGGACCAACCCACATCTGTG 1049
Db      622 spGluGlyIuTrpProItrPrlValSerIeuHisAlaIleuGlnIyHisIleCysG 642
QY      1050 GAGGCGAGCTGATGAGCGCCAGGAGGCTCATGCGGCCACATGGCTTCTCGTGAACC 1109
Db      642 IylAlaSerIeuIleSerProAsnIthrPleuValSerAlaAlaHisCysIyrIleAspAsp 662
QY      1110 GGGAGAAAGTCTCGAGAGC-----TGGAAGTGTACGCG 1145
Db      662 rg-----GlyPheArgIyrSerAspProThrGlnIthrPheAlaPheIeuG 677
QY      1146 GCACACGAACTGTGACACGATGCTGAGCAGCTCC-----A 1184
Db      677 Iy-----LeuHisAspGlnSerGlnArgSerAlaProGlyValGlnIuArgIyl 694
QY      1185 TTGCGGAGATCATGATCAACAGCAATTTCACCGATGAGAGGAGGACGATGACATCGGCC 1244

```

```

Db      694 eulysArgIleIleSerHisProPhePheAsnAspPheThrPheAspIyrAspIleAla 714
QY      1245 TCATGCGGCTGTCCAAACCCCTGACCTGTCCGCTACATCCACCTGCTTGCCTCCCA 1304
Db      714 eulGluIeuGluIysProAlaGluIyrIyrSerSerMetValArgProIleCysIeuProA 734
QY      1305 TGCATGAGCAGACTTTAGCTCAATGAGCTGCTGATGATCACAGGCTTGGCAAGACA 1364
Db      734 spAlaSerHisValPheProAlaGlyValAlaIleTrpValIthrPrlHisThr- 753
QY      1365 GGGAGACAGATGACAAAGACATCCCTTCTCCGGAGGTGACAGTCAATCTCATGACT 1424
Db      754 --GlnIyrGlyIuThrGlnIyAlaIleuIleuGlnIysGlyIuIleArgValIleAsn 773
QY      1425 TCAAGAAATGCATGACTACTTGTCTATGACAGTTACCTTACCCCAAGATGATGTGTG 1484
Db      773 IthrThrCysGluAsnIleu-----ProIuGlnIleThrProArgMetMetCysv 791
QY      1485 CTGGGACCTTCTGGGGGAGAGACTCTGCCAGAGGACAGCGGGGGCTCTT---G 1541
Db      791 alGlyPheIeuSerGlyIyAlaIysPercysGlnIyAspSerGlyIyProIeuSers 811
QY      1542 TCTGTGACAGAAACAACCGCTGTGACTCTGACAGGTGTCAACAGCTGGGACAGCTGTG 1601
Db      811 ervalGluAlaAspGlyArgIlePheGlnIlaGlyValIalSerTrpGlyAspGlyCysA 831
QY      1602 GCCAGAGAAACAACCTGTGTGTATCACCAAGTGCAGAGAAATTCTTCCTGATT 1657
Db      831 lagInArgAsnIysProGlyValIyrThrArgIeuProIeuPheArgAspIrlle 849

RESULT 12
US-11-254-185-6
: Sequence 6, Application US/11254185
: Publication No. US20060099625A1
: GENERAL INFORMATION:
: APPLICANT: CRAIK, CHARLES S.
: APPLICANT: TAKEUCHI, TOSHIOHKO
: APPLICANT: SCHUMAN, MARC
: TITLE OF INVENTION: MT-SPI SERINE PROTEASE
: FILE REFERENCE: 28644-701.302
: CURRENT APPLICATION NUMBER: US/11/254,185
: CURRENT FILING DATE: 2005-10-18
: NUMBER OF SEQ ID NOS: 83
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 6
: LENGTH: 255
: TYPE: PRT
: ORGANISM: Artificial
: FEATURE:
: OTHER INFORMATION: Protein fragment/domain
US-11-254-185-6

Alignment Scores:
Pred. No.: 5,04e-24 Length: 255
Score: 525.50 Matches: 105
Percent Similarity: 57.7% Conservative: 38
Best Local Similarity: 42.3% Mismatch: 88
Query Match: 15.8% Indels: 17
DB: Gaps: 6

US-10-806-370-11 (1-1748) x US-11-254-185-6 (1-255)
QY      971 ATCGTGGAGGGGGCGCTGCGCTGAGTACGAAGTGGCTTTGGCAAGTGAAGTTCACATTC 1030
Db      1 IlevAlGlyIyArgAspThrIserIeuGlyArgItrProItrPrlValIserIeuArgIyr 20
QY      1031 GGCACCAACCAATCTGTGGAGGACGCTATTGACGCCAGTGGGTGCTCATGCGGCC 1090
Db      21 AspGlyAlaHisIeuCysGlyIySerIeuIeuSerCylAspItrPrlValIeuHisAla 40
QY      1091 CACTGCTTCTTGTGACCCGCGAGAAAGCTCTGAGGCTGGAAGGCTGACGCGGACCC 1150
Db      41 HisCys---PheProIuLysAsnArgValIeuSerArgItrPargValPheAlaGlyAla 59

```


Pred. No.: 5.04e-24 Length: 255
 Score: 525.50 Matches: 105
 Percent Similarity: 57.7% Conservative: 38
 Best Local Similarity: 42.3% Mismatches: 88
 Query Match: 15.8% Indels: 17
 DB: 7 Gaps: 6

US-10-806-370-11 (1-1748) x US-11-253-869-6 (1-255)

QY 971 ATGCTGGAGAGGCGCTGGCGCTGCGATAGCAAGTGGCGTGGCAAGTGTGCACTTC 1030
 Db 1 IlevaGlygIyArGaSPThrSerLeuGlyArGTrProTrpGlnValSerLeuArGTr 20
 QY 1031 GGCACCAACCCATCTGTGAGGACCGCTCATTAACCCAGTGGGGCTCACGCGGCC 1090
 Db 21 AspGlyAlaHisLeuCySGlyGlySerLeuSerGlyAspTrpValLeuThrAlaAla 40
 QY 1091 CACTGCTTCTTCGTGACCGGAGAGAGTCTGTGAGGCGTGGCAAGTGTACGCGGAC 1150
 Db 41 HisCys---PheProGlnuArGaSPThrSerLeuSerArGTrPArGValPheAlaGlyAla 59
 QY 1151 -----AGCAACCTGCACAGTTCCTGAGGACCTTCATTCGCGAGATCATC 1201
 Db 60 ValAlaGlnAlaSerProHisGlyLeu-----GlnLeuGlyValGlnAlaValAlaTyr 77
 QY 1202 AACAGCAATTAC-----ACCATAGAGAGAGACGACTATGATGATGCC 1243
 Db 78 HisGlyGlyTyrLeuProPheArGaSPThrProAsnSerGlnuAsnSerAsnAspIleAla 97
 QY 1244 CTGATGGGCTGTGTCAGACCCCTGACCTGCTGCATATCCACCTGCTGGCTGCC 1303
 Db 98 LeuValHisLeuSerSerProLeuProLeuThrGlnuTyrTrpGlnProValCysLeuPro 117
 QY 1304 ATGCATGACAGACCTTACCTTCATGAGACCTGCTGATCAGAGCTTTGGCAAGC 1363
 Db 118 AlAlaGlyGlnAlaLeuValAspGlyLysIleCysThrValThrGlyTyrGlyAsnThr 137
 QY 1364 AGGAGACAGATGACAGACATCCCTCTCCGAGAGAGTGCAGGTCAATCTCATGAC 1423
 Db 138 GlnTyrTyrGlnGlnAlaGly---ValLeuGlnGlnuAlaArGValProIleLeSer 156
 QY 1424 TTCAGAAATGCAATGACTTCTGTATGACAGCTTACCTTACCCAGAGATGATGT 1483
 Db 157 AsnAspValCysAsnGlyAlaAspPheTyrGlnuGlnuIleLysProLysMetPheCys 176
 QY 1484 GCTGGGACCTTCGTGGGAGAGACTCTGCGAGAGAGACAGCGGGGCGCTCTTGT 1543
 Db 177 AlaGlyTyrProGlnuGlyIleAspAlaCysGlnuGlnuAspSerGlyProPheVal 196
 QY 1544 TGTGAG-----CAGAACACCGCTGTGACTGTGAGAGTGTACACAGCTGGGGC 1591
 Db 197 CysGlnuAspSerIleSerArGTrProArGTrPArGLeuCySGlyIleValSerTrpGly 216
 QY 1592 ACAGGCTGTGGCCAGAGAAACAACTGTGTGTACCAAGAGAGCAAGATCTTCCC 1651
 Db 217 ThrGlyCysAlaLeuAlaGlnuLysProGlyValTyrThrLysValSerAspPheArGln 236
 QY 1652 TGGATTACAGCAAGATGAGAGC 1675
 Db 237 TriPleHeGlnAlaIleLysThr 244

NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 38
 ; LENGTH: 255
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-253-869-38

Alignment Scores:
 Pred. No.: 5.04e-24 Length: 255
 Score: 525.50 Matches: 105
 Percent Similarity: 57.7% Conservative: 38
 Best Local Similarity: 42.3% Mismatches: 88
 Query Match: 15.8% Indels: 17
 DB: 7 Gaps: 6

US-10-806-370-11 (1-1748) x US-11-253-869-38 (1-255)

QY 971 ATGCTGGAGAGGCGCTGGCGCTGCGATAGCAAGTGGCGTGGCAAGTGTGCACTTC 1030
 Db 1 IlevaGlygIyArGaSPThrSerLeuGlyArGTrProTrpGlnValSerLeuArGTr 20
 QY 1031 GGCACCAACCCATCTGTGAGGACCGCTCATTAACCCAGTGGGGCTCACGCGGCC 1090
 Db 21 AspGlyAlaHisLeuCySGlyGlySerLeuSerGlyAspTrpValLeuThrAlaAla 40
 QY 1091 CACTGCTTCTTCGTGACCGGAGAGAGTCTGTGAGGCGTGGCAAGTGTACGCGGAC 1150
 Db 41 HisCys---PheProGlnuArGaSPThrSerLeuSerArGTrPArGValPheAlaGlyAla 59
 QY 1151 -----AGCAACCTGCACAGTTCCTGAGGACCTTCATTCGCGAGATCATC 1201
 Db 60 ValAlaGlnAlaSerProHisGlyLeu-----GlnLeuGlyValGlnAlaValAlaTyr 77
 QY 1202 AACAGCAATTAC-----ACCATAGAGAGAGACGACTATGATGATGCC 1243
 Db 78 HisGlyGlyTyrLeuProPheArGaSPThrProAsnSerGlnuAsnSerAsnAspIleAla 97
 QY 1244 CTGATGGGCTGTGTCAGACCCCTGACCTGCTGCATATCCACCTGCTGGCTGCC 1303
 Db 98 LeuValHisLeuSerSerProLeuProLeuThrGlnuTyrTrpGlnProValCysLeuPro 117
 QY 1304 ATGCATGACAGACCTTACCTTCATGAGACCTGCTGATCAGAGCTTTGGCAAGC 1363
 Db 118 AlAlaGlyGlnAlaLeuValAspGlyLysIleCysThrValThrGlyTyrGlyAsnThr 137
 QY 1364 AGGAGACAGATGACAGACATCCCTCTCCGAGAGAGTGCAGGTCAATCTCATGAC 1423
 Db 138 GlnTyrTyrGlnGlnAlaGly---ValLeuGlnGlnuAlaArGValProIleLeSer 156
 QY 1424 TTCAGAAATGCAATGACTTCTGTATGACAGCTTACCTTACCCAGAGATGATGT 1483
 Db 157 AsnAspValCysAsnGlyAlaAspPheTyrGlnuGlnuIleLysProLysMetPheCys 176
 QY 1484 GCTGGGACCTTCGTGGGAGAGACTCTGCGAGAGAGACAGCGGGGCGCTCTTGT 1543
 Db 177 AlaGlyTyrProGlnuGlyIleAspAlaCysGlnuGlnuAspSerGlyProPheVal 196
 QY 1544 TGTGAG-----CAGAACACCGCTGTGACTGTGAGAGTGTACACAGCTGGGGC 1591
 Db 197 CysGlnuAspSerIleSerArGTrProArGTrPArGLeuCySGlyIleValSerTrpGly 216
 QY 1592 ACAGGCTGTGGCCAGAGAAACAACTGTGTGTACCAAGAGAGCAAGATCTTCCC 1651
 Db 217 ThrGlyCysAlaLeuAlaGlnuLysProGlyValTyrThrLysValSerAspPheArGln 236
 QY 1652 TGGATTACAGCAAGATGAGAGC 1675
 Db 237 TriPleHeGlnAlaIleLysThr 244

Search completed: September 16, 2006, 02:48:37
 Job time : 76 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:04:25 ; Search time 199 Seconds
(without alignments)
1291.234 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999
Sequence: 1 MERSHONASPARFPASAGS.....TEVLPWITYSKMSEVFRPKS 562

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_8: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *
9: Geneseqp2005s: *
10: Geneseqp2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2999 | 100.0 | 562 | 5 | AAE17238 Human tra |
| 2 | 2999 | 100.0 | 569 | 8 | ADH17440 Human NOV |
| 3 | 2991 | 99.7 | 562 | 4 | AAE01943 Human tra |
| 4 | 2991 | 99.7 | 562 | 7 | AD110393 Human cel |
| 5 | 2991 | 99.7 | 562 | 8 | AD146917 Human tra |
| 6 | 2991 | 99.7 | 562 | 9 | ADY50146 Human end |
| 7 | 2963 | 98.8 | 688 | 4 | AAE01944 Human tra |
| 8 | 2963 | 98.8 | 688 | 7 | AD110395 Human cel |
| 9 | 2963 | 98.8 | 688 | 8 | AD146919 Human tra |
| 10 | 2963 | 98.8 | 688 | 7 | ADY50054 Human end |
| 11 | 2963 | 98.6 | 581 | 7 | ADY50054 Human end |
| 12 | 2948.5 | 98.3 | 586 | 8 | ADH17414 Human NOV |
| 13 | 2948.5 | 98.3 | 586 | 8 | ADH17434 Human NOV |
| 14 | 2948.5 | 98.3 | 593 | 8 | ADH17430 Human NOV |
| 15 | 2943.5 | 98.1 | 586 | 8 | ADH17452 Human NOV |
| 16 | 2942.5 | 98.1 | 586 | 8 | ADH17450 Human NOV |
| 17 | 2942 | 98.1 | 581 | 5 | ADZ75552 Human mos |
| 18 | 2775.5 | 92.5 | 537 | 5 | AAU82746 Amino aci |
| 19 | 2775.5 | 92.5 | 537 | 8 | ADH17438 Human NOV |
| 20 | 2737.5 | 91.3 | 556 | 9 | AAE20240 Novel hum |
| 21 | 2686 | 89.6 | 542 | 8 | ADH17432 Human NOV |
| 22 | 2414.5 | 80.5 | 471 | 8 | ABM83371 Human dia |
| 23 | 2291.5 | 76.4 | 477 | 6 | ABP56842 Human den |

| | | | | | |
|----|--------|------|-----|---|--------------------|
| 24 | 2287 | 76.3 | 421 | 4 | AAB85042 Human SER |
| 25 | 2287 | 76.3 | 421 | 8 | ADH17442 Human NOV |
| 26 | 2254 | 75.2 | 419 | 7 | ADE29367 Human ser |
| 27 | 2219.5 | 74.0 | 486 | 5 | AAE18171 Human pro |
| 28 | 2207 | 73.6 | 491 | 4 | AAB93442 Human pro |
| 29 | 2130.5 | 71.0 | 446 | 6 | ADH17454 Human NOV |
| 30 | 2108 | 70.3 | 446 | 6 | ADP56843 Human den |
| 31 | 2097 | 69.9 | 382 | 8 | ADH17444 Human NOV |
| 32 | 2059 | 68.7 | 406 | 8 | ADH17436 Human NOV |
| 33 | 2059 | 68.7 | 406 | 8 | ADH17448 Human NOV |
| 34 | 2026 | 67.6 | 401 | 8 | ADH17422 Human NOV |
| 35 | 2026 | 67.6 | 418 | 7 | ADE79008 Human pro |
| 36 | 2016 | 67.2 | 401 | 8 | ADH17426 Human pro |
| 37 | 2013.5 | 67.1 | 449 | 7 | ADE78981 Human pro |
| 38 | 2012 | 67.1 | 401 | 8 | ADH17428 Human NOV |
| 39 | 2008 | 67.0 | 451 | 7 | ADE78982 Human pro |
| 40 | 1906.5 | 63.6 | 375 | 7 | ADE78993 Human pro |
| 41 | 1842.5 | 61.4 | 370 | 8 | ADH17424 Human NOV |
| 42 | 1839.5 | 61.3 | 367 | 8 | ADH17420 Human NOV |
| 43 | 1311 | 43.7 | 296 | 3 | AAV72108 Human ser |
| 44 | 1311 | 43.7 | 372 | 3 | AAV72092 Human ser |
| 45 | 1311 | 43.7 | 372 | 5 | ABG30789 Human ser |

ALIGNMENTS

RESULT 1
ID AAE17238 standard; protein; 562 AA.
AC AAE17238;
XX
DT 18-APR-2002 (first entry)
DE Human transmembrane serine protease.
XX
KW Human; transmembrane serine protease; gene therapy; metastasis; tumour;
KW chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;
KW atherosclerosis; neurodegenerative disease; neuroprotective; cytoskeletal;
KW pathogenic infection; antiinflammatory; antiarteriosclerotic;
KW antibacterial.
XX
OS Homo sapiens.
XX
PN WO200196538-A2.
XX
PD 20-DEC-2001.
XX
PF 12-JUN-2001; 2001WO-EP006618.
XX
PR 13-JUN-2000; 2000US-0211224P.
PR 13-APR-2001; 2001US-0283353P.
PR 16-APR-2001; 2001US-0283648P.
XX
PA (FARB) BAYER AG.
XX
PI Xiao Y, Gedrich R;
XX
DR WPI; 2002-098065/13.
DR N-PSDB; AAD27734.
XX
PT Novel isolated polynucleotide encoding transmembrane serine protease
PT polypeptide, for treating chronic obstructive pulmonary disease, tumor
PT angiogenesis, inflammation, atherosclerosis and neurodegenerative
PT disease.
XX
PS Claim 1; Fig 1; 120bp; English.
XX
CC The present invention relates to an isolated polynucleotide encoding a
CC transmembrane serine protease polypeptide. Transmembrane serine protease
CC gene is useful in gene therapy. The invention also relates to a
CC pharmaceutical composition which is useful for modulating the activity of

transmembrane serine protease in a disease, such as chronic obstructive pulmonary disease (COPD), metastasis of malignant cells, tumour angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease or pathogenic infection. Transmembrane serine protease is useful as a bait protein in a two-hybrid or three-hybrid assay. The polypeptide is useful for generating antibodies against it and in various assay systems. The present sequence is a human transmembrane serine protease

Sequence 562 AA;

Query Match 100.0%; Score 2999; DB 5; Length 562;

Best Local Similarity 100.0%; Pred. No. 8,2e-189; Mismatches 0; Indels 0; Gaps 0;

Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MERDGHNASPARTSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
1 MERDGHNASPARTSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
121 PVGAAPIRSSPARSPATRATESPCTSLPKTMEGQKOLPLICVLLILALVSLITL 180
121 PVGAAPIRSSPARSPATRATESPCTSLPKTMEGQKOLPLICVLLILALVSLITL 180
121 PVGAAPIRSSPARSPATRATESPCTSLPKTMEGQKOLPLICVLLILALVSLITL 180
181 FQFMGHTGIRKXKRESCPKAVACDGVDCLEKSDLEGCYRPFMDLSLKITGSSHQ 240
181 FQFMGHTGIRKXKRESCPKAVACDGVDCLEKSDLEGCYRPFMDLSLKITGSSHQ 240
241 WLPICSSWNDSYSKTCQQLGPGSAHRTTEVAHDFPANSFSLRYNSTIOGSLHRSBP 300
241 WLPICSSWNDSYSKTCQQLGPGSAHRTTEVAHDFPANSFSLRYNSTIOGSLHRSBP 300
301 SORVSLSCSHCGRAMTGRIVGALASDSKPMQVSHFQTHICGTLIDAOQVLTAA 360
301 SORVSLSCSHCGRAMTGRIVGALASDSKPMQVSHFQTHICGTLIDAOQVLTAA 360
301 SORVSLSCSHCGRAMTGRIVGALASDSKPMQVSHFQTHICGTLIDAOQVLTAA 360
361 HCFVTRREKVLGKMYVAGTSMHLQPEASIAEIIINSNTYDEBDDVIALMRLEKPLT 420
361 HCFVTRREKVLGKMYVAGTSMHLQPEASIAEIIINSNTYDEBDDVIALMRLEKPLT 420
421 LSAHHPACLPWHGQTFSLNETCWTGFKTRTETDOKTSPFLREVOVNLIDPKKNDYLV 480
421 LSAHHPACLPWHGQTFSLNETCWTGFKTRTETDOKTSPFLREVOVNLIDPKKNDYLV 480
421 LSAHHPACLPWHGQTFSLNETCWTGFKTRTETDOKTSPFLREVOVNLIDPKKNDYLV 480
481 YDSYLTTPRMWCAGLRGGRDSCGDSGGLVCEQNNRWYLAGVTSWGTGCGRRKPGVYT 540
481 YDSYLTTPRMWCAGLRGGRDSCGDSGGLVCEQNNRWYLAGVTSWGTGCGRRKPGVYT 540
541 KYTEVLPMIYSKMESEVFRKS 562
541 KYTEVLPMIYSKMESEVFRKS 562

RESULT 2

ADH17440
ADH17440 standard; protein; 569 AA.

ADH17440;

11-MAR-2004 (first entry)

Human NOV12n protein - SEQ ID 130.

NOVX; antidiabetic; anorectic; cardiac; hypotensive;
antiartherosclerotic; anorectic; vinuclide; antibacterial; fungicide;
proliferative; neuroprotective; antiparkinsonian;
anticoronary; osteopathic; antiarthritic; antiinflammatory;
dermatological; antidiabetic; antidiabetic; metabolic; diabetes;
obesity; infection; anorexia; cancer; cardiovascular; hypertension;
atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
epilepsy; immune; osteoarthritis; haemopoietic;

inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
cell differentiation; proliferation; haemopoiesis; wound healing;
angiogenesis; gene therapy; chromosome mapping; tissue typing;
pharmacogenomic; human.

Homo sapiens.

WO2003093432-A2.

13-NOV-2003.

02-MAY-2003; 2003WO-US013690.

02-MAY-2002; 2002US-0377321P.
08-MAY-2002; 2002US-0378730P.
29-MAY-2002; 2002US-0383075P.
29-MAY-2002; 2002US-0384044P.
30-MAY-2002; 2002US-0384215P.
30-MAY-2002; 2002US-0384215P.
30-MAY-2002; 2002US-0384296P.
30-MAY-2002; 2002US-0384327P.
30-MAY-2002; 2002US-0384352P.
31-MAY-2002; 2002US-0385211P.
02-JUL-2002; 2002US-0393331P.
09-AUG-2002; 2002US-0402154P.
09-AUG-2002; 2002US-0402171P.
09-AUG-2002; 2002US-0402204P.
09-AUG-2002; 2002US-0402205P.
22-AUG-2002; 2002US-0405175P.
27-AUG-2002; 2002US-0406125P.
23-SEP-2002; 2002US-0412954P.
30-SEP-2002; 2002US-0414975P.
07-OCT-2002; 2002US-0416661P.
24-OCT-2002; 2002US-0420851P.
31-OCT-2002; 2002US-0422547P.
01-MAY-2003; 2003US-00428275.

(CURA-) CURAGEN CORP.

Alvarez E, Anderson DM, Boldog FL, Catterton E, Edinger SR,
Fernandes ER, Gerlach VL, Gorman L, Grosse WM, Guo X, Ji W,
Kerkda R, Li L, Macdougall JR, Padigaru M, Patuwajan M,
Peterson JD, Raetelli L, Shinkets RA, Spytek KA, Stone DJ,
Vernet CAM, Voss EZ, Zhong M;
MPI; 2004-053040/05.
N-PSDB; ADH17439.

New isolated NOVX polypeptide, useful for preventing, diagnosing or
treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

Claim 1; SEQ ID NO 130; 478bp; English.

The invention relates to a novel isolated NOVX polypeptide. The
polypeptide of the invention demonstrates antidiabetic, anorectic,
cardiac, hypotensive, antiatherosclerotic, anorectic, vinuclide,
antibacterial, fungicide, procoronary, neurotropic, neuroprotective,
antiparkinsonian, anticoronary, osteopathic, antiarthritic,
antiinflammatory, dermatological, antidiabetic and antiproliferative
activities. The polypeptide, nucleic acid molecules and antibodies may
be useful in the manufacture of a medicament for treating metabolic
disorders, diabetes, obesity, infectious diseases (viral, bacterial,
fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
diseases including hypertension and atherosclerosis, neurodegenerative
disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
disorders such as osteoarthritis, haemopoietic disorders, inflammatory
skin disorders, asthma and various types of dyslipidaemia. The nucleic
acids and polypeptides may also be used as targets for the identification
of small molecules that modulate or inhibit neurogenesis, cell
differentiation, cell proliferation, haemopoiesis, wound healing and
angiogenesis, in gene therapy and the in generation of antibodies that
bind immunospecifically to NOVX substances for use in therapeutic or


```
Db 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Qy 61 PACTPPGRASPGARSPAQASPARASPALASLSRSSSGSSARSASVTTSPTRYLVYRAT 120
Db 61 PACTPPGRASPGARSPAQASPARASPALASLSRSSSGSSARSASVTTSPTRYLVYRAT 120
Qy 121 PVGAVPIRSSPARSAPATRAITRESPTSLPKFTWREGQKQPLIGCVLLIALVLSLIL 180
Db 121 PVGAVPIRSSPARSAPATRAITRESPTSLPKFTWREGQKQPLIGCVLLIALVLSLIL 180
Qy 181 FQFMQGHGIRYKQRRSCPHAVRCQVVOCKLSDELGCVRPMDKSLKITYSGSSHQ 240
Db 181 FQFMQGHGIRYKQRRSCPHAVRCQVVOCKLSDELGCVRPMDKSLKITYSGSSHQ 240
Qy 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSRCP 300
Db 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSRCP 300
Qy 301 SQRYVLSQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
Db 301 SQRYVLSQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
Qy 361 HCFVTRREKVLKGVNVAAGTNSLHQLPBAASIAEIIINSNTDEDDVDIALMLSKPLT 420
Db 361 HCFVTRREKVLKGVNVAAGTNSLHQLPBAASIAEIIINSNTDEDDVDIALMLSKPLT 420
Qy 421 LSAHIHPACLPMMHGTFSLNCTCWTGFKTRTDDKTSPLREYQVNLIDFKKCNLYLV 480
Db 421 LSAHIHPACLPMMHGTFSLNCTCWTGFKTRTDDKTSPLREYQVNLIDFKKCNLYLV 480
Qy 481 YDSYLTFRMWCAGDLRGGRDSCQDGGPLVCBQNNRWYLAGVTSWGTGCGQRNKPQVYT 540
Db 481 YDSYLTFRMWCAGDLRGGRDSCQDGGPLVCBQNNRWYLAGVTSWGTGCGQRNKPQVYT 540
Qy 541 KYTEVLPWYSKMSSEVFRKFS 562
Db 541 KYTEVLPWYSKMSSEVFRKFS 562

RESULT 4
AD110393
ID AD110393 standard; proteain; 562 AA.
AC AD110393;
XX
XX 22-APR-2004 (first entry)
DT
DB Human cell surface protease #12.
XX
XX therapeutic agent; plasmin; protease specific antigen; PSA;
KW cell-surface protease-associated disease; cancer; ocular disease;
KW cardiovascular disease; chronic inflammatory disease; wound;
KW circulatory disorder; dermatological disorder; rheumatoid arthritis;
KW porriasis; diabetic retinopathy; pterygium;
KW excimer laser surgery scarring; glaucoma filtering surgery scarring;
KW macular degeneration; crest syndrome; solid neoplasm; vascular tumor;
KW melanoma; Kaposi's sarcoma; human; cell surface protease.
XX
XX Homo sapiens.
OS
PN WO200295007-A2.
XX
XX 28-NOV-2002.
PD
XX 23-MAY-2002; 2002MO-US016819.
PF
XX 23-MAY-2001; 2001US-0293267P.
PR
XX (CORV-) CORVAS INT INC.
PA
XX Madison EL, Sempke JE, Vlausk GP, Kemp ST, Komandla M, Stew DY,
PI
XX MPI; 2003-221280/21.
DR
```

```
DR N-PSDB; AD110392.
XX
XX Novel conjugate useful for treating cell-surface protease-associated
PT disease, comprises a therapeutic agent and a peptidic or nucleic acid
PT substrate linked to it optionally by a peptidic linker.
XX
XX Claim 9; SEQ ID NO 24; 581bp; English.
PS
XX The invention comprises a conjugate that consists of a therapeutic agent
CC and a peptide substrate (optionally linked via linker). The peptide
CC substrate is proteolytically cleaved by a cell surface protease or a
CC soluble, released or shed form of it, to liberate the therapeutic agent,
CC the conjugate of the invention is not substantially cleaved by plasmin or
CC protease specific antigen (PSA). The conjugate of the invention is useful
CC for treating a cell-surface protease-associated disease such as: cancer,
CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
CC wounds, circulatory disorders, dermatological disorders, rheumatoid
CC arthritis, porriasis, diabetic retinopathy, recurrence of pterygium,
CC scarring from excimer laser surgery, scarring from glaucoma filtering
CC surgery, macular degeneration, crest syndrome, solid neoplasm, vascular
CC tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence
CC represents a human cell surface protease.
XX
XX Sequence 562 AA:
SQ
Query Match 99.7%; Score 2991; DB 7; Length 562;
Best Local Similarity 99.8%; Pred. No. 2.8e-188;
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Qy 61 PACTPPGRASPGARSPAQASPARASPALASLSRSSSGSSARSASVTTSPTRYLVYRAT 120
Db 61 PACTPPGRASPGARSPAQASPARASPALASLSRSSSGSSARSASVTTSPTRYLVYRAT 120
Qy 121 PVGAVPIRSSPARSAPATRAITRESPTSLPKFTWREGQKQPLIGCVLLIALVLSLIL 180
Db 121 PVGAVPIRSSPARSAPATRAITRESPTSLPKFTWREGQKQPLIGCVLLIALVLSLIL 180
Qy 181 FQFMQGHGIRYKQRRSCPHAVRCQVVOCKLSDELGCVRPMDKSLKITYSGSSHQ 240
Db 181 FQFMQGHGIRYKQRRSCPHAVRCQVVOCKLSDELGCVRPMDKSLKITYSGSSHQ 240
Qy 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSRCP 300
Db 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSRCP 300
Qy 301 SQRYVLSQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
Db 301 SQRYVLSQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
Qy 361 HCFVTRREKVLKGVNVAAGTNSLHQLPBAASIAEIIINSNTDEDDVDIALMLSKPLT 420
Db 361 HCFVTRREKVLKGVNVAAGTNSLHQLPBAASIAEIIINSNTDEDDVDIALMLSKPLT 420
Qy 421 LSAHIHPACLPMMHGTFSLNCTCWTGFKTRTDDKTSPLREYQVNLIDFKKCNLYLV 480
Db 421 LSAHIHPACLPMMHGTFSLNCTCWTGFKTRTDDKTSPLREYQVNLIDFKKCNLYLV 480
Qy 481 YDSYLTFRMWCAGDLRGGRDSCQDGGPLVCBQNNRWYLAGVTSWGTGCGQRNKPQVYT 540
Db 481 YDSYLTFRMWCAGDLRGGRDSCQDGGPLVCBQNNRWYLAGVTSWGTGCGQRNKPQVYT 540
Qy 541 KYTEVLPWYSKMSSEVFRKFS 562
Db 541 KYTEVLPWYSKMSSEVFRKFS 562

RESULT 5
ADJ46917
ID ADJ46917 standard; proteain; 562 AA.
```

XX AC AD46917;
XX DT 06-MAY-2004 (first entry)
XX DE Human transmembrane serine protease (MTSP)-related polypeptide #2.
XX
XX Human; transmembrane serine protease; MTSP; cell surface protease;
XX plasmin; prostate specific antigen; PSA; proliferative disease;
XX cell-surface protease-associated disease; autoimmune disease;
XX inflammatory disease; infectious disease; endocrine disease; cancer;
XX ocular disorder; cardiovascular disorder; chronic inflammatory disease;
XX wound; circulatory disorder; dermatological disorder; restenosis;
XX rheumatoid arthritis; psoriasis; diabetic retinopathy;
XX laser surgery scarring; glaucoma filtering surgery scarring;
XX macular degeneration; CREST syndrome; bacterial infection; viral disease;
XX solid neoplasm; vascular tumour; lung; colon; prostate; melanoma;
XX Kaposi's sarcoma; enzyme.
XX
XX Homo sapiens.
XX US2004001801-A1.
XX 01-JAN-2004.
XX 23-MAY-2002; 2002US-00156214.
XX 23-MAY-2002; 2002US-00156214.
XX (CORV-) CORVAS INT INC.
XX Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Stew DV;
XX WPI; 2004-190126/18.
XX N-PSDB; ADJ46916.
XX
XX Conjugate useful for treating e.g. cancer, cell-surface protease-
XX associated diseases, comprising a peptidic substrate or nucleic acid
XX substrate linked to a therapeutic agent through a linker.
XX
XX Claim 9; SEQ ID NO 24; 361pp; English.
XX
XX The invention relates to a conjugate comprising a therapeutic agent and a
XX peptidic substrate or nucleic acid substrate linked to the agent
XX optionally through a linker or peptidic linker, where the peptidic
XX substrate is proteolytically cleaved by a cell surface protease or a
XX soluble, released or shed form conjugate to liberate the agent and the
XX conjugate is not substantially cleaved by plasmin or prostate specific
XX antigen (PSA). The conjugate is useful for treating a disease, which
XX involves administering a conjugate to a subject, where the disease is
XX preferably a proliferative disease or a cell-surface protease-associated
XX disease. The diseases include autoimmune diseases, inflammatory diseases,
XX infectious diseases and endocrine diseases. The conjugate is useful for
XX treating a cell-surface protease-associated disease, which involves
XX administering a conjugate comprising an agent and a peptidic substrate to
XX a subject exhibiting symptoms of a cell-surface protease-associated
XX disorder, where the disease is selected from cancer, ocular disorders,
XX cardiovascular disorders, chronic inflammatory diseases, wounds,
XX circulatory disorders, dermatological disorders, restenosis, rheumatoid
XX arthritis, psoriasis, diabetic retinopathies, scarring from laser
XX surgery, scarring from glaucoma filtering surgery, macular degeneration,
XX CREST syndrome, bacterial infections, viral diseases, solid neoplasms and
XX vascular tumours such as lung cancer, colon cancer, prostate cancer,
XX melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
XX serine protease (MTSP)-related polypeptide of the invention.
XX
XX Sequence 562 AA;
XX
XX Query Match 99.7%; Score 2991; DB 8; Length 562;
XX Best Local Similarity 99.8%; Pred. No. 2.8e-188;
XX Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MERDGHNASPARTSPAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
QY 61 PACTPPGRASPGRASPAGASPAQASPALASISRSSSGSSARSASVTTSTRTYLVNAT 120
DB 61 PACTPPGRASPGRASPAGASPAQASPALASISRSSSGSSARSASVTTSTRTYLVNAT 120
QY 121 PVGAVPIRSSPARAPATRATRESPGTSLPKFTREBQKOLPLIGCVLLTALVSLIIL 180
DB 121 PVGAVPIRSSPARAPATRATRESPGTSLPKFTREBQKOLPLIGCVLLTALVSLIIL 180
QY 181 FQFMQHTGIRKYNQRESCEPAHVRCDGVNDCUKSDELGCVRPDMKSLTKIYSGSSHQ 240
DB 181 FQFMQHTGIRKYNQRESCEPAHVRCDGVNDCUKSDELGCVRPDMKSLTKIYSGSSHQ 240
QY 241 WLPICSSNMWDSYEKTCQQLGFESAHRTTEVARNDFANSFSLIRYNSTIOESLHRSCEP 300
DB 241 WLPICSSNMWDSYEKTCQQLGFESAHRTTEVARNDFANSFSLIRYNSTIOESLHRSCEP 300
QY 301 SORYSISQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHCGSTLIDAQWVLTAA 360
DB 301 SORYSISQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHCGSTLIDAQWVLTAA 360
QY 361 HCFVTRKVLGKMYAGTISNLHQLPEASIAEIIINSNTDEBDYDIALMRLSKPLT 420
DB 361 HCFVTRKVLGKMYAGTISNLHQLPEASIAEIIINSNTDEBDYDIALMRLSKPLT 420
QY 421 LSAHIHPACLBPMQGTSLNETCWTITGFKTRFTDCKTSPFLREVQVNLIPFKCNDIYV 480
DB 421 LSAHIHPACLBPMQGTSLNETCWTITGFKTRFTDCKTSPFLREVQVNLIPFKCNDIYV 480
QY 481 YDSVLTTPRMCAGLRGGRDSCGDSGGLVCEQNNRYLAGVTSWGCGGGGRNKPQYTT 540
DB 481 YDSVLTTPRMCAGLRGGRDSCGDSGGLVCEQNNRYLAGVTSWGCGGGGRNKPQYTT 540
QY 541 KYTEVLPMWYSKMESEVFRKS 562
DB 541 KYTEVLPMWYSKMESEVFRKS 562
RESULT 6
ADY50146
ID ADY50146 standard; protein, 562 AA.
AC ADY50146;
XX 19-MAY-2005 (first entry)
DE Human endotheliase-2L.
XX
XX Antiarthritic; Antirheumatic; Antipsoriatic; Antidiabetic;
XX Ophthalmological; Cardiovascular-Gen.; Vulnary; Antiinflammatory;
XX Vasootropic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
XX rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
XX dermatological disease; cancer; neoplasm; endotheliase-2; enzyme.
XX
XX Homo sapiens.
XX WO2005019270-A2.
XX 03-MAR-2005.
XX 12-AUG-2004; 2004WO-US026148.
XX 14-AUG-2003; 2003US-0495005P.
XX 14-NOV-2003; 2003US-0520164P.
XX
XX (DYAX-) DYAX CORP.
XX (DENND-) DENNDEN CORP.
XX Madison EL, Nixon A;
XX WPI; 2005-202609/21.

DR N-PSDB; ADY50145.

| PT | Novel protein capable of inhibiting endotheliase-2, useful for treating or preventing angiogenesis related disorder e.g. cancer. |
|-----|--|
| 786 | |

PS Disclosure; SEQ ID NO 94; 157pp; English.

CC The invention relates to an isolated protein (I) which comprises a heavy
CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)

immunoglobulin variable domain sequence, where the first and second immunoglobulin variable domain sequences from an antigen binding site that specifically binds to human endothelialase-2. (I) is useful for detecting an endothelialase or endothelialase activity in a sample, for modulating an activity of an E12-expressing cell, for modulating proteolysis, for killing or inhibiting growth of a cell, for detecting endothelialase in a subject, for modulating endothelialase activity in a subject, for treating or preventing a disorder characterized by unwanted angiogenesis in a subject. The disorder is chosen from rheumatoid arthritis, psoriasis, diabetic retinopathies, ocular disorder such as pterygial recurrence, scarring excimer laser surgery and glaucoma filtering surgery, cardiovascular disorders, chronic inflammatory disorders, wound repair, circulatory disorders, crest syndromes, dermatological disorders and cancers. The present sequence represents the human endothelialase-2S.

cc human endothel1ase-25.

disorders Raynaud's phenomenon, crest syndromes such as calcinosis, oesophageal, dyomeolysis, sclerodactyly and teanglectasis, dermatological disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum, vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome and Osler-Weber-Rendu syndrome and ocular disorders such as blindness caused by ocular neovascular disease, corneal graft neovascularisation, macular degeneration, retinopathy of prematurity, retrolental fibroplasia and corneal neovascularisation. The nucleic acids of the invention are also used in gene therapy. The invention also provides method for screening compounds that modulate angiogenesis

Sequence 688 AA:

Query Match 98.8%; Score 2963; DB 4; Length 688;
Best Local Similarity 99.3%; Pred. No. 2.4e-186;

Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 MERDSHGNSAPARTSPASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
1 MERDSHGNSAPARTSPASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
181 FOFMOGHGIRYKQRESCEPKAVRCGVNDCKLSDELGCVRFPMDSLLKITYGSSHQ 240
181 FOFMOGHGIRYKQRESCEPKAVRCGVNDCKLSDELGCVRFPMDSLLKITYGSSHQ 240
181 FOFMOGHGIRYKQRESCEPKAVRCGVNDCKLSDELGCVRFPMDSLLKITYGSSHQ 240
241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSCEP 300
241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSCEP 300
241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSCEP 300
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360
361 HCFVTRKEVLEGMKVYAGTSLNHLPEASIAEIIINSNTYDEEDVDIALMRLSKPLT 420
361 HCFVTRKEVLEGMKVYAGTSLNHLPEASIAEIIINSNTYDEEDVDIALMRLSKPLT 420
361 HCFVTRKEVLEGMKVYAGTSLNHLPEASIAEIIINSNTYDEEDVDIALMRLSKPLT 420
421 LSAHIHPACLPWAGTFSINETCWTGFGKTRTETDCKTSPFLREVQVNLIDPKKCNLYLV 480
421 LSAHIHPACLPWAGTFSINETCWTGFGKTRTETDCKTSPFLREVQVNLIDPKKCNLYLV 480
421 LSAHIHPACLPWAGTFSINETCWTGFGKTRTETDCKTSPFLREVQVNLIDPKKCNLYLV 480
481 YDSYITTPMMACAGDLRGGRDSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGGRNRPGYTT 540
481 YDSYITTPMMACAGDLRGGRDSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGGRNRPGYTT 540
481 YDSYITTPMMACAGDLRGGRDSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGGRNRPGYTT 540
541 KYTEVLPWISKMESEVR 558
541 KYTEVLPWISKMESEVR 558
541 KYTEVLPWISKMESEVR 558

RESULT 8

AD110395
AD110395 standard; protein; 688 AA.

AD110395;
AD110395;

22-APR-2004 (first entry)

Human cell surface protease #13.

therapeutic agent; plasmin; protease specific antigen; PSA;
cell-surface protease-associated disease; cancer; ocular disease;
cardiovascular disease; chronic inflammatory disorder; rheumatoid arthritis;
circulatory disorder; dermatological disorder; rheumatoid arthritis;

psoriasis; diabetic retinopathy; pterygium;
excimer laser surgery scarring; glaucoma filtering surgery scarring;
macular degeneration; crest syndrome; solid neoplasm; vascular tumour;
melanoma; Kaposi's sarcoma; human; cell surface protease.

Homo sapiens.

WO200295007-A2.

28-NOV-2002.

23-MAY-2002; 2002WO-US016819.

23-MAY-2001; 2001US-0293267P.

(CORV-) CORVAS INT INC.

Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV,

WPI; 2003-221280/21.

N-PSDB; AD110394.

Novel conjugate useful for treating cell-surface protease-associated disease, comprises a therapeutic agent and a peptidic or nucleic acid substrate linked to it optionally by a peptidic linker.

Claim 9; SEQ ID NO 26; 581bp; English.

The invention comprises a conjugate that consists of a therapeutic agent and a peptide substrate (optionally linked via linker). The peptide substrate is proteolytically cleaved by a cell surface protease or a soluble, released or shed form of it, to liberate the therapeutic agent, the conjugate of the invention is not substantially cleaved by plasmin or protease specific antigen (PSA). The conjugate of the invention is useful for treating a cell-surface protease-associated disease such as: cancer, ocular diseases, cardiovascular diseases, chronic inflammatory diseases, wounds, circulatory disorders, dermatological disorders, rheumatoid arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium, scarring from excimer laser surgery, scarring from glaucoma filtering surgery, macular degeneration, crest syndromes, solid neoplasms, vascular tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence represents a human cell surface protease.

Sequence 688 AA:

Query Match 98.8%; Score 2963; DB 7; Length 688;
Best Local Similarity 99.3%; Pred. No. 2.4e-186;

Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 MERDSHGNSAPARTSPASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
1 MERDSHGNSAPARTSPASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWEGQKQLPGICVLLIALVSLIIL 180
181 FOFMOGHGIRYKQRESCEPKAVRCGVNDCKLSDELGCVRFPMDSLLKITYGSSHQ 240
181 FOFMOGHGIRYKQRESCEPKAVRCGVNDCKLSDELGCVRFPMDSLLKITYGSSHQ 240
181 FOFMOGHGIRYKQRESCEPKAVRCGVNDCKLSDELGCVRFPMDSLLKITYGSSHQ 240
241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSCEP 300
241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSCEP 300
241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSCEP 300
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360
301 SORITSLDCHSCGRLAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLLDAQWVLTAA 360

| | | | |
|----------|---|---|-----|
| Qy | 361 | HCFEFTREVLVAGMKVYKVAATSNLHOLPEAASIAETIINSNYDEEDDYIALMRLSKPLT | 420 |
| Db | 361 | HCFEFTREVLVAGMKVYKVAATSNLHOLPEAASIAETIINSNYDEEDDYIALMRLSKPLT | 420 |
| Qy | 421 | LSAHIHPACLPMMHGQTFSLINETCMWITGFGKTRFTDRTKTSPPFLREVOVNIIDPKKCNDIYV | 480 |
| Db | 421 | LSAHIHPACLPMMHGQTFSLINETCMWITGFGKTRFTDRTKTSPPFLREVOVNIIDPKKCNDIYV | 480 |
| Qy | 481 | YDSYLTTPPMKCGDRIKGGDSCQSGGGLVCEONNRWLVIAGTWSGTCGQGNKRCVYT | 540 |
| Db | 481 | YDSYLTTPPMKCGDRIKGGDSCQSGGGLVCEONNRWLVIAGTWSGTCGQGNKRCVYT | 540 |
| Qy | 541 | KVTEVLPMIYSKMESEVR | 558 |
| Db | 541 | KVTEVLPMIYSKMESEVR | 558 |
| RESULT 9 | | | |
| ID | ADJ46919 | standard; protein; 688 AA. | |
| AC | ADJ46919; | | |
| XX | 06-MAY-2004 | (first entry) | |
| DT | | | |
| DE | Human transmembrane serine protease (MTSP)-related polypeptide #3. | | |
| XX | Human; transmembrane serine protease; MTSP; cell surface protease; | | |
| KM | plasma; prostate specific antigen; PSA; proliferative disease; | | |
| KM | cell-surface protease-associated disease; autoimmune disease; | | |
| KM | inflammatory disease; infectious disease; endocrine disease; cancer; | | |
| KM | ocular disorder; cardiovascular disorder; chronic inflammatory disease; | | |
| KM | wound; circulatory disorder; dermatological disorder; restenosis; | | |
| KM | rheumatoid arthritis; psoriasis; diabetic retinopathy; | | |
| KM | laser surgery scarring; glaucoma; filtering surgery scarring; | | |
| KM | macular degeneration; CREST syndrome; bacterial infection; viral disease; | | |
| KM | solid neoplasm; vascular tumour; lung; colon; prostate; melanoma; | | |
| KM | Kaposi's sarcoma; enzyme. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PN | US2004001801-A1. | | |
| PD | 01-JAN-2004. | | |
| XX | | | |
| PF | 23-MAY-2002; 2002US-00156214. | | |
| XX | | | |
| PR | 23-MAY-2002; 2002US-00156214. | | |
| XX | | | |
| PA | (CORV-) CORVAS INT INC. | | |
| PI | Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV; | | |
| XX | | | |
| DR | WPI; 2004-190126/18. | | |
| XX | | | |
| XX | N-PSDB; ADJ46918. | | |
| PT | Conjugate useful for treating e.g. cancer, cell-surface protease- | | |
| PT | associated diseases, comprising a peptidic substrate or nucleic acid | | |
| PT | substrate linked to a therapeutic agent through a linker. | | |
| XX | | | |
| PS | Claim 9; SEQ ID NO 26; 361pp; English. | | |
| XX | | | |
| CC | The invention relates to a conjugate comprising a therapeutic agent and a | | |
| CC | peptidic substrate or nucleic acid substrate linked to the agent | | |
| CC | optionally through a linker or peptidic linker, where the peptidic | | |
| CC | substrate is proteolytically cleaved by a cell surface protease or a | | |
| CC | soluble, released or shed form conjugate to liberate the agent and the | | |
| CC | conjugate is not substantially cleaved by plasmin or prostate specific | | |
| CC | antigen (PSA). The conjugate is useful for treating a disease, which | | |
| CC | involves administering a conjugate to a subject, where the disease is | | |
| CC | preferably a proliferative disease or a cell-surface protease-associated | | |
| CC | disease. The diseases include autoimmune diseases, inflammatory diseases, | | |
| CC | infectious diseases and endocrine diseases. The conjugate is useful for | | |

| | | |
|----|---|---|
| | CC | treating a cell-surface protease-associated disease, which involves |
| | CC | administering a conjugate comprising an agent and a peptidic substrate to |
| | CC | a subject exhibiting symptoms of a cell-surface protease-associated |
| | CC | disease, where the disease is selected from cancer, ocular disorders, |
| | CC | cardiovascular disorders, chronic inflammatory diseases, wounds, |
| | CC | circulatory disorders, dermatological disorders, restenosis, rheumatoid |
| | CC | arthritis, psoriasis, diabetic retinopathies, scarring from laser |
| | CC | surgery, scarring from glaucoma filtering surgery, macular degeneration, |
| | CC | CREST syndrome, bacterial infections, viral diseases, solid neoplasms and |
| | CC | vascular tumours such as lung cancer, colon cancer, prostate cancer, |
| | CC | melanoma and Kaposi's sarcoma. This sequence represents a transmembrane |
| | CC | serine protease (MSP)-related polypeptide of the invention. |
| | XX | |
| SQ | | Sequence 688 AA: |
| | | |
| | Query Match | 98.8%; Score 2963; DB 8; Length 688; |
| | Best Local Similarity | 99.3%; Pred. No. 2,4e-186; |
| | Matches | 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0; |
| OY | | |
| Dd | 1 MERDSDHGASDPATPPSAGASPAAQSPAGTPPGRGASPAQASPAQSPAGTTPGGASPAQAS 1 MERDSHGNASSAPRTIPPSAGSPAQASPAQTTPGRGASPAQASPAQSPAGTTPGGASPAQAS | 60 |
| OY | 61 PACGPPEPGASCGRASPAQASPARPASALASISRSSSSGRSSARSAVTTSTPTRYLVNAT | 120 |
| Dd | 61 PACGPPPGASCGRASPAQASPARPASALASISRSSSGRSSARSASVTTSPTRVLYVNAT | 120 |
| OY | 121 PVGAVPTRSSPARAPATRAPRARESPGSLPKFTRREGOKOPLIGCVLLILALVLSLITL | 180 |
| Dd | 121 PVGAVPIRRSSPARAPATRAPRESPTGLPKFTREGEKOLPLIGCVLLILLVLVSLLITL | 180 |
| OY | 181 FQFMQGHTGIIRYKEQRSCPHNAVRCDGVNDCKLSDELGCVRPDMDXLLIKIYGSSHQ | 240 |
| Dd | 181 FQMVGHTGIIRYKKQRSCPHGARICGVNDCKKSDELCVRFPDMDSLIIKIYGSSSHQ | 240 |
| OY | 241 WLPICSSNMNDYSSEKTQQQLGFESAHRTTEVAHRDFANSFSILIRYNSTIOESLHRSCEP | 300 |
| Dd | 241 WLPICSSNMNDYSSEKTCQQQLGFESAHRTEVARHPDFSNSILIRYNSTIOESIHSRECP | 300 |
| OY | 301 SQRIYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFQTHICQGITLLIDAQMVLTPAA | 360 |
| Dd | 301 SQRIYSIQCSHCGRAMTAGRIVGALASDSKMPQVSLHFQTHICQGITLLIDAQMVLTPAA | 360 |
| OY | 361 HCFPVTEREKVLBGKNKYAAGTSNLHLPRASIABIINSNTDBEDVDIALMLSKRLT | 420 |
| Dd | 361 HCFFVTERKVLEBKVKYAAGTSNLHLPRASAIAIINSNTDBEDEDVDIALMLSKRLT | 420 |
| OY | 421 LSAIHHPACLPBMHQOTSLNETCMTITGKGKRERETDDKTSPLRLAEYOVLIDFKKCNDYLIV | 480 |
| Dd | 421 LSAIHHPACLPBMHQOTSINETCWITTFGKTRERETDDKTSPLRLREVQNLIDPFKCNLDYLIV | 480 |
| OY | 481 YDSYLTFRPMCAGLRGGRDSCGDSDGGLVCENNNRWYLAGVTSWGTGCGQRNKPGVYT | 540 |
| Dd | 481 YDSYLTFRPMCAGLRGGRDSCGDSDGGLVCENNNRWYLAGVTSWGTCGQRNKPGVYT | 540 |
| OY | 541 KYTEVLPMVIYSKOMSEVR 558 | |
| Dd | 541 KYTEVLPMVIYSKOMENRAQ 558 | |
| | | |
| | RESULT 10 | |
| | ID ADY50054 | |
| | ADY50054 standard; protein; 688 AA. | |
| XX | | |
| AC | ADY50054; | |
| XX | | |
| DT | 19-MAY-2005 (first entry) | |
| XX | | |
| XX | Human endotheliasae-2L. | |
| KM | Anthraxitidis; Antirheumatic; Antiportariatis; Antidiabetic; | |
| KM | Ophthalmological; Cardiovascular-Gen.; Vulnerary; Antiinflammatory; | |
| KM | Vasotrophic; Dermatologic; Cytostatic; Angiangiogenic; angiogenesis; | |

| QY | Query Match | 98.8% | Score 2363 | DB 9 | Length 688 |
|----|---|----------------|--------------------|----------|------------|
| QY | Best Local Similarity | 99.3% | Pred. No. 2.4e-186 | | |
| QY | Matches 554 | Conservative 2 | Mismatches 2 | Indels 0 | Gaps 0 |
| QY | 1 MERDSHGASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS | 60 | | | |
| DB | 1 MERDSHGASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS | 60 | | | |
| QY | 61 PACGPPGASPERASPAQASPARASPALASLRSSSGSSSRASSVTTPSPRYLVART | 120 | | | |
| DB | 61 PACGPPGASPERASPAQASPARASPALASLRSSSGSSSRASSVTTPSPRYLVART | 120 | | | |
| QY | 121 PVGAVPIRSSPARASPARATRESGTSLPKRTWEGQKQLPLICVLLILAVLSLIL | 180 | | | |
| DB | 121 PVGAVPIRSSPARASPARATRESGTSLPKRTWEGQKQLPLICVLLILAVLSLIL | 180 | | | |
| QY | 121 PVGAVPIRSSPARASPARATRESGTSLPKRTWEGQKQLPLICVLLILAVLSLIL | 180 | | | |
| DB | 121 PVGAVPIRSSPARASPARATRESGTSLPKRTWEGQKQLPLICVLLILAVLSLIL | 180 | | | |
| QY | 181 FQFWGHTGIRYKQRESCPKAIVACDGVDCUKLSDELGCYRFPMDXSLKITYGSSHQ | 240 | | | |
| DB | 181 FQFWGHTGIRYKQRESCPKAIVACDGVDCUKLSDELGCYRFPMDXSLKITYGSSHQ | 240 | | | |
| QY | 241 WLPICSSMNDBYSEKTCQQLAFESAHRTTEVAHDFANSFILRNSITQESLHRSECP | 300 | | | |
| DB | 241 WLPICSSMNDBYSEKTCQQLAFESAHRTTEVAHDFANSFILRNSITQESLHRSECP | 300 | | | |
| QY | 301 SGRYISLQCSHGCLAMTGRIVGALASDCKMPQVSLHFGTTHICGGLIDAOWLTAA | 360 | | | |

| | | | |
|-----------|---|---|-----|
| Db | 301 | SQRYSLSQCSHCGLAAMTGRIVGALALASDKKPMQVSLHFGTHHICGGTLIDAQWULTAA | 360 |
| Qy | 361 | HCFPTREKVELGKWKVAVGATSNLHQLPEAASIAEIIINSNYTDEEDDYDIALMRLSKPLT | 420 |
| Db | 361 | HCFPTREKVELGKWKVAVGATSNLHQLPEAASIAEIIINSNYTDEEDDYDIALMRLSKPLT | 420 |
| Qy | 421 | LSAHHHPACLPMHGGTFSLSNETCWIITGPKTETEDTKTSPFLREVQVNLIDPKKCDYLV | 480 |
| Db | 421 | LSAHHHPACLPMHGGTFSLSNETCWIITGPKTETEDTKTSPFLREVQVNLIDPKKCDYLV | 480 |
| Qy | 481 | YDSYLTPPMWCGADLRGRDSCQSGSGPLVCEONNRWYLAGVTSWGTSCGGRNKPQVYT | 540 |
| Db | 481 | YDSYLTPPMWCGADLRGRDSCQSGSGPLVCEONNRWYLAGVTSWGTSCGGRNKPQVYT | 540 |
| Qy | 541 | KYTEVLPWYISKMESEVR | 558 |
| Db | 541 | KYTEVLPWYISKMESEVR | 558 |
| RESULT 11 | | | |
| ID | ADE31743 | AD31743 standard; protein; 581 AA. | |
| XX | | | |
| AC | ADE31743; | | |
| DT | 29-JAN-2004 | (first entry) | |
| XX | | | |
| DE | Human 2914 | protein #SEQ ID 100. | |
| XX | | | |
| KW | Antiartherosclerotic; cardiact; vasotropic; antiinflammatory; | | |
| KW | thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy; | | |
| KW | cardiovascular; disorder; ischaemia; aortic bending; | | |
| KW | vascular heart disease; endocarditis; atrial fibrillation; heart failure; | | |
| KW | angina; cardiomyopathy; cardiac death. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PN | WO2003065984-A2. | | |
| XX | | | |
| PD | 14-AUG-2003. | | |
| XX | | | |
| PF | 29-JAN-2003; 2003WO-US002571. | | |
| XX | | | |
| PR | 01-FEB-2002; 2002US-0353224P. | | |
| PR | 15-MAR-2002; 2002US-0364529P. | | |
| PR | 19-APR-2002; 2002US-0373861P. | | |
| PR | 29-APR-2002; 2002US-0376287P. | | |
| PR | 12-JUN-2002; 2002US-0386080P. | | |
| PR | 24-JUN-2002; 2002US-0390971P. | | |
| PR | 03-JUL-2002; 2002US-0394130P. | | |
| PR | 10-JUL-2002; 2002US-0394797P. | | |
| PR | 21-AUG-2002; 2002US-0404904P. | | |
| PR | 23-AUG-2002; 2002US-0405450P. | | |
| PR | 04-SEP-2002; 2002US-0408070P. | | |
| PR | 06-NOV-2002; 2002US-0424300P. | | |
| PR | 05-DEC-2002; 2002US-0431042P. | | |
| PR | 05-DEC-2002; 2002US-0431079P. | | |
| XX | | | |
| PA | (MILL-) MILLENNIUM PHARM INC. | | |
| XX | | | |
| PI | Logan TV, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M; | | |
| PI | Stagliano N, Petrodin J, Rodrigue-Way A; | | |
| XX | | | |
| DR | WPI; 2003-731468/69. | | |
| XX | N-PSDB; ADE31742. | | |
| XX | | | |
| PT | Identifying a compound capable of treating a cardiovascular disorder | | |
| PT | (e.g. atherosclerosis) comprises assaying the ability of the compound to | | |
| PT | modulate the expression or activity of e.g. 1682, 6169 or 6193 | | |
| PT | polypeptide or nucleic acid. | | |
| XX | | | |
| PS | Disclosure; SEQ ID NO 100; 328bp; English. | | |

XX The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates antidiabetic, anorectic,
 CC cardiact, hypotensive, antiarteriosclerotic, anorectic, virucide,
 CC antibacterial, fungicide, protozoacide, neurotropic, neuroprotective,
 CC antiparkinsonian, anticonvulsant, osteopathic, antiinflammatory,
 CC antiinflammatory, dermatological, antiasmatic and antihypertensive
 CC activities. The polypeptide, nucleic acid molecules and antibodies may
 CC be useful in the manufacture of a medicament for treating metabolic
 CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
 CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
 CC diseases including hypertension and atherosclerosis, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
 CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
 CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
 CC acids and polypeptides may also be used as targets for the identification
 CC of small molecules that modulate or inhibit neurogenesis, cell
 CC differentiation, cell proliferation, haemopoiesis, wound healing and
 CC angiogenesis, in gene therapy and the in generation of antibodies that
 CC bind immunospecifically to NOVX substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids may be further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine and pharmacogenomics. The current sequence is that of the human
 CC NOVX protein of the invention.

SO Sequence 586 AA:

Query Match 98.3%; Score 2948.5; DB 8; Length 586;
 Best Local Similarity 99.1%; Pred. No. 1.8e-185;
 Matches 554; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MERDSHGASAPRPPASAPASAPGTPGRASAPASAPASAPGTPGRASAPAS 60
 DB 1 MERDSHGASAPRPPASAPASAPGTPGRASAPASAPASAPGTPGRASAPAS 60
 QY 61 PACTPPGRASPCR-----ASPAQASPARASPALASLSSSSSSGRSSASASAVTTSPTRYV 115
 DB 61 PACTPPGRASPCRASPAQASPARASPALASLSSSSSSGRSSASASAVTTSPTRYV 120
 QY 116 LVKATPPGAVPIRSSPARSAPATATRESPTSLPKFTWREGOKOLPIIGCVLLIALIV 175
 DB 121 LVKATPPGAVPIRSSPARSAPATATRESPTSLPKFTWREGOKOLPIIGCVLLIALIV 180
 QY 176 SLTLFQWOGHTGIRYKRESCPKAVRCDGVDCGLKSDDELGCYAFDMDKSLTKYS 235
 DB 181 SLTLFQWOGHTGIRYKRESCPKAVRCDGVDCGLKSDDELGCYAFDMDKSLTKYS 240
 QY 236 GSSHQWLPICSSNNNDYSSEKTCQOLGFSARHRTTEVAHRDPFANSFILRYNSTIOESLH 295
 DB 241 GSSHQWLPICSSNNNDYSSEKTCQOLGFSARHRTTEVAHRDPFANSFILRYNSTIOESLH 300
 QY 296 RSECPSPORYSISLQCSHCGLRMTGRIVGALASDSKMPVOVSLHFGTHICGTLIDAQW 355
 DB 301 RSECPSPORYSISLQCSHCGLRMTGRIVGALASDSKMPVOVSLHFGTHICGTLIDAQW 360
 QY 356 VLTAAHCFVTVREKVLBEKMYAGTNSNHLPEASTAIEIINSYDDEEDVDIALMRL 415
 DB 361 VLTAAHCFVTVREKVLBEKMYAGTNSNHLPEASTAIEIINSYDDEEDVDIALMRL 420
 QY 416 SKPLTLNHIHACLPNMGQTFSLNETGWTGFGKTRTDDTSPFLREVOVNLIDPFKC 475
 DB 421 SKPLTLNHIHACLPNMGQTFSLNETGWTGFGKTRTDDTSPFLREVOVNLIDPFKC 480
 QY 476 NDYLVYDSYLTLPMMACADLRGRSDSCGDSGGPLVCEQNNRWYLAGVTSNGTGGGKNK 535
 DB 481 NDYLVYDSYLTLPMMACADLRGRSDSCGDSGGPLVCEQNNRWYLAGVTSNGTGGGKNK 540
 QY 536 PGVYTKVTEVLPWTYSKME 554
 DB 541 PGVYTKVTEVLPWTYSKME 559

RESULT 13

ADH17434
 ID ADH17434 standard; protein; 586 AA.
 XX
 AC ADH17434;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human NOV12k protein - SEQ ID 124.
 XX
 KW NOVX; antidiabetic; anorectic; cardiact; hypotensive;
 KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
 KW protozoacide; neurotropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
 KW dermatological; antiasmatic; antihypertensive; metabolic; diabetes;
 KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
 KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
 KW epilepsy; immune; osteoarthritis; haemopoietic;
 KW inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
 KW cell differentiation; proliferation; haemopoiesis; wound healing;
 KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
 KW pharmacogenomic; human.
 XX
 OS Homo sapiens.
 XX
 FN WO2003093432-A2.
 XX
 PD 13-NOV-2003.
 XX
 PF 02-MAY-2003; 2003WO-US013690.
 XX
 PR 02-MAY-2002; 2002US-0377321P.
 PR 08-MAY-2002; 2002US-0378730P.
 PR 24-MAY-2002; 2002US-0383075P.
 PR 29-MAY-2002; 2002US-0384044P.
 PR 30-MAY-2002; 2002US-0384215P.
 PR 30-MAY-2002; 2002US-0384296P.
 PR 30-MAY-2002; 2002US-0384327P.
 PR 30-MAY-2002; 2002US-0384352P.
 PR 31-MAY-2002; 2002US-0385211P.
 PR 02-JUL-2002; 2002US-0393333P.
 PR 09-AUG-2002; 2002US-0402154P.
 PR 09-AUG-2002; 2002US-0402171P.
 PR 09-AUG-2002; 2002US-0402204P.
 PR 09-AUG-2002; 2002US-0402205P.
 PR 22-AUG-2002; 2002US-0405175P.
 PR 27-AUG-2002; 2002US-0406129P.
 PR 23-SEP-2002; 2002US-0412954P.
 PR 30-SEP-2002; 2002US-0414975P.
 PR 07-OCT-2002; 2002US-0416661P.
 PR 24-OCT-2002; 2002US-0420851P.
 PR 31-OCT-2002; 2002US-0422547P.
 PR 01-MAY-2003; 2003US-00428275.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alvarez E, Anderson DM, Boldog FL, Catterton E, Edinger SR,
 PI Fernandes ER, Gerlach VL, Gorman L, Grosse WM, Guo X, Ji W,
 PI Kekuda R, Li L, Macdougall JR, Padigaru M, Pattnarajan M,
 PI Peterson JD, Rastelli L, Shinkets RA, Szytek KA, Stone DJ,
 PI Vernet CM, Voss EZ, Zhong M;
 XX
 DR WPI: 2004-053040/05.
 DR N-PSDB: ADH17433.
 XX
 PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
 PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
 PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
 XX
 PS Claim 1; SEQ ID NO 124; 478bp; English.
 XX
 CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates antidiabetic, anorectic,

CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritis,
CC antiinflammatory, dermatological, antiepileptic,
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.
CC
XX

SQ Sequence 586 AA;

Query Match 98.1%; Score 2943.5; DB 8; Length 586;
Best Local Similarity 98.9%; Pred. No. 3.9e-185;
Matches 553; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 MERDSHGNSAPARTSPASAPASPAQSPAGPPGRASPAPASPAQSPAGPPGRASPAPAS 60
DB 1 MERDSHGNSAPARTSPASAPASPAQSPAGPPGRASPAPASPAQSPAGPPGRASPAPAS 60
QY 61 PAGTPPGRASPGR-----ASPAQASPARASPALASISRSSSGRSSSARSASVTTSPTRY 115
DB 61 PAGTPPGRASPGRASPAQSPAPASPARASPALASISRSSSGRSSSARSASVTTSPTRY 120
QY 116 LVRATPVGAVPRTSSPARASAPATRTATRSPTGSLPKFTWRBQKQPLIGCVLLIALVY 175
DB 121 LVRATPVGAVPRTSSPARASAPATRTATRSPTGSLPKFTWRBQKQPLIGCVLLIALVY 180
QY 176 SLIIIFQFQGHGTGIRYKQRESCPHAYRCGVVDCCLKSDCLGCVFPMDSKSLIKIYS 235
DB 181 SLIIIFQFQGHGTGIRYKQRESCPHAYRCGVVDCCLKSDCLGCVFPMDSKSLIKIYS 240
QY 236 GSSHQWLPICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPFANSFILRYNSTIOESLH 295
DB 241 GSSHQWLPICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPFANSFILRYNSTIOESLH 300
QY 296 RSECP8QRYISIQCSHCGLRANTGRIVGALASDSKWPQVSLHFGTTHICGTLIDAW 355
DB 301 RSECP8QRYISIQCSHCGLRANTGRIVGALASDSKWPQVSLHFGTTHICGTLIDAW 360
QY 366 VLTAAHCFVTRTEKYLEGKVKVAGTSNLHQLPEASIAIITNSNYTDEEDYDIALMRL 415
DB 361 VLTAAHCFVTRTEKYLEGKVKVAGTSNLHQLPEASIAIITNSNYTDEEDYDIALMRL 420
QY 416 SKPLTLSAHIPACLPMGQGTFSLNETCMITGFKTRTDDKTSPLREVOYNLIDFKKC 475
DB 421 SKPLTLSAHIPACLPMGQGTFSLNETCMITGFKTRTDDKTSPLREVOYNLIDFKKC 480
QY 476 NDYLYVDSYLTFRNMCCAGDLRGRDSCQDSDSGPLVCBONNRWYLAGTWSGTGCGQRNK 535
DB 481 NDYLYVDSYLTFRNMCCAGDLRGRDSCQDSDSGPLVCBONNRWYLAGTWSGTGCGQRNK 540
QY 536 PGVYTKVTEVLPMIYSKME 554
DB 541 PGVYTKVTEVLPMIYSKME 559

Search completed: September 16, 2006, 03:08:13
Job time : 201 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:08:31; Search time 43 Seconds

(without alignments)
1257.531 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999

Sequence: 1 MERSHGNASAPRTPSAGAS.....TEVLPWIKMESEVRPRKS 562

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 611.5 | 20.4 | 417 | S00845 | hepsin (EC 3.4.21. |
| 2 | 592.5 | 19.8 | 416 | S33777 | hepsin (EC 3.4.21. |
| 3 | 575 | 19.2 | 1019 | A56318 | enteropeptidase (E |
| 4 | 574 | 19.1 | 638 | 1 KQMSPL | plasma kallikrein |
| 5 | 572.5 | 19.0 | 638 | 1 KQMSPL | plasma kallikrein |
| 6 | 568.5 | 18.9 | 855 | 2 JG7731 | membrane-bound arg |
| 7 | 565.5 | 18.5 | 638 | 1 KQRTPL | plasma kallikrein |
| 8 | 554 | 18.5 | 1034 | 1 A53663 | enteropeptidase (E |
| 9 | 549 | 18.3 | 1113 | 2 JEO315 | low-density lipopr |
| 10 | 548.5 | 18.3 | 1035 | 1 A43090 | enteropeptidase (E |
| 11 | 536 | 17.9 | 761 | 2 JCS759 | brain-specific ser |
| 12 | 532 | 17.7 | 625 | 1 KFHU1 | coagulation factor |
| 13 | 514 | 17.1 | 812 | 1 PLEBO | plasma (EC 3.4.21 |
| 14 | 510.5 | 17.0 | 343 | 1 A57014 | proctasin (EC 3.4. |
| 15 | 508 | 16.9 | 274 | 2 JCA171 | tryptase (EC 3.4.2 |
| 16 | 506.5 | 16.9 | 455 | 2 A61545 | plasma (EC 3.4.21 |
| 17 | 505 | 16.8 | 276 | 2 A38654 | mast cell proteina |
| 18 | 495 | 16.5 | 460 | 2 B61545 | plasma (EC 3.4.21 |
| 19 | 495 | 16.5 | 812 | 1 PLMS | plasma (EC 3.4.21 |
| 20 | 493.5 | 16.5 | 810 | 1 PLHU | plasma (EC 3.4.21 |
| 21 | 491 | 16.4 | 270 | 2 S56160 | mast cell tryptase |
| 22 | 488.5 | 16.3 | 273 | 2 A47246 | tryptase (EC 3.4.2 |
| 23 | 480 | 16.0 | 275 | 2 C35863 | serine proteinase |
| 24 | 480 | 16.0 | 786 | 1 A47547 | plasma (EC 3.4.21 |
| 25 | 480 | 16.0 | 790 | 1 PLRG | plasma (EC 3.4.21 |
| 26 | 479 | 16.0 | 275 | 2 A35863 | tryptase (EC 3.4.2 |
| 27 | 479 | 16.0 | 415 | 1 A34170 | actosin (EC 3.4.21 |
| 28 | 478 | 15.9 | 275 | 2 B35863 | tryptase (EC 3.4.2 |
| 29 | 477.5 | 15.9 | 418 | 2 A37344 | actosin (EC 3.4.21 |

| | | | | | |
|----|-------|------|------|----------|--------------------|
| 30 | 477 | 15.9 | 275 | 2 A32410 | tryptase (EC 3.4.2 |
| 31 | 473 | 15.8 | 437 | 2 S18407 | actosin (EC 3.4.21 |
| 32 | 472.5 | 15.8 | 810 | 2 B30848 | plasma (EC 3.4.21 |
| 33 | 471.5 | 15.7 | 4548 | 1 S00657 | apoptotain(a) (EC |
| 34 | 470 | 15.7 | 436 | 2 UX0172 | actosin (EC 3.4.21 |
| 35 | 468.5 | 15.6 | 1524 | 2 T30337 | polyprotein - Altr |
| 36 | 466.5 | 15.6 | 1420 | 2 A32869 | apolipoprotein(a) |
| 37 | 466 | 15.5 | 421 | 2 S29599 | actosin (EC 3.4.21 |
| 38 | 466 | 15.5 | 655 | 1 A46688 | hepatocyte growth |
| 39 | 457.5 | 15.3 | 274 | 2 A45754 | tryptase (EC 3.4.2 |
| 40 | 457 | 15.2 | 431 | 2 S47538 | actosin (EC 3.4.21 |
| 41 | 456 | 15.2 | 810 | 2 I46260 | plasma (EC 3.4.21 |
| 42 | 453.5 | 15.1 | 421 | 1 S11674 | actosin (EC 3.4.21 |
| 43 | 447 | 14.9 | 434 | 1 A35005 | u-plasminogen acti |
| 44 | 438 | 14.6 | 237 | 2 S68702 | tryptase (EC 3.4.2 |
| 45 | 435 | 14.5 | 562 | 1 UKHUT | t-plasminogen acti |

ALIGNMENTS

RESULT 1

S00845
hepsin (EC 3.4.21.-) - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Oct-2004
C/Accession: S00845
R/Entry: S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.
Biochemistry 27, 1067-1074, 1988
A/Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane dome
A/Reference number: S00845; MUID:88209431; PMID:2835076
A/Accession: S00845
A/Molecule type: mRNA
A/Residues: 1-417 <LE>
A/Cross-References: UNIPROT:P05981; UNIPARC:UPI000003FE67; EMBL:X07732; NID:G32063; PTDN:
C/Genetics:
A/Gene: GDB:HPN; TMPSR1; hepsin
A/Cross-References: GDB:135685; OMIM:142440
A/Map position: 19q11-19q13.2
C/Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F.23-45/Domain: transmembrane #status predicted <TM>
F.163-400/Domain: trypsin homology <TRY>
F.188-204,291-359,322-338,349-381/Dissulfide bonds: #status predicted
F.203,257,353/Active site: His, Asp, Ser #status predicted

Query Match 20.4%; Score 611.5; DB 1; Length 417;
Best Local Similarity 36.0%; Pred. No. 2.3e-26;
Matches 134; Conservative 62; Mismatches 147; Indels 29; Gaps 9;

| | | | |
|----|-----|--|-----|
| QY | 209 | VVDCKLSGDE--LGCVRFDWIKSLIKIYSGSSHOWLPICSSNMWDSYSEKTCQQLGFESA | 266 |
| DB | 39 | IVAVLTKNSDQEPFLVQVSSADARLMDVDTKBTGTRLLCSSNSNRNVAGLSEENGFLLRA | 98 |
| QY | 267 | HRTTEVARD-----FANSPILRYNSTIOESLHRSBCPSQRYVSLQCSHGLRAM | 317 |
| DB | 99 | LTHSELVTRTGANSTGFCVDEBGRLPHTQRLLEVISVCCPRRFLAIIQDCGRRL | 158 |
| QY | 318 | TGRIVGALASDSKMPQVSLHFGTHICGGLTIDAQVLTAAFCFVTRKYLEGMKV | 376 |
| DB | 159 | PVDRIYVGRDSTLSGMPQVSLRYDGAHICGSLISGDMVLTAAHC--PPERRVLSRRRV | 217 |
| QY | 377 | YAGT---SNLHQLPEAASIAEIIINSY-----TDEDDVDIALMRSLKPLTSAHIHP | 427 |
| DB | 218 | PAGAVQAQSPHGL--QLGVQAVVHGGILPRDPNSSENSDIALVHSSPLPLEYIOP | 275 |
| QY | 428 | ACLPMHGGTFSINMETCWTGFGKTRTDKTSPLREVOVNLIDFKKNDVLYVDSYLTLP | 487 |
| DB | 276 | VCLPAAQALVDGKICTVGTGNGNTQYQGQAG-VLQEARVPLISNDVCGADFYNGQIKP | 334 |
| QY | 488 | RMCMAGDLRGDRDSQGGSGPVLCE-----QNNRYVLAGVTSWGCQORNRKPGVYTVT | 543 |
| DB | 335 | KMFAGYEGGIDACQGSQGGSPVCEDSISRTPRRLCGIVSGTGCALAQPGVYTVKS | 394 |

Oy 544 EULPMYKMS 555
Db 395 DFERWIFQAIKT 406

RESULT 2

hepsin (BC 3.4.21.-) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004
C:Accession: S33777; S32013
R:Barley, D.; Raymond, F.; Nick, H.
Biochim. Biophys. Acta 1173, 350-352, 1993
A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.
A:Reference number: S33777; MUID:93305733; PMID:8318546
A:Accession: S33777
A:Structure: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <FAR>
A:Cross-references: UNIPROT:Q05511; UNIPARC:UPI00004BCD9; EMBL:X70900; NID:957928; PIDN:
F:12-44/Domain: hydrolase; liver; serine proteinase; transmembrane protein
F:162-389/Domain: trypsin homology <TRY>
F:187-203,280-358,321-337,348-380/Dissulfide bonds: #status predicted
F:202,256,352/Active site: His, Asp, Ser #status predicted

Query Match 19.8%; Score 592.5; DB 1; Length 416;

Best Local Similarity 36.3%; Pred. No. 2,56-25;
Matches 138; Conservative 59; Mismatches 138; Indels 45; Gaps 12;

Oy 209 VVDCKLKDE--LGCVRPMDKSLIKIYSGSSHOWLPICSSNMNDSYEKTCQQLGFPSA 266
Db 38 IVTILRSDOBYLVQVQLSPGSRLLVLDKTEGTRLLCSSRSNARVAGLCSEMGFLRA 97
Oy 267 HRTTEVARDF-----NSFS-----ILKNSTIOESLHRSCEQSORYISIQCHC 312
Db 98 -----LHASELDVRTAGNAGTSGFPCVDEGGPLAQLRLDVISVCDGRGPRFLATCCDC 152
Oy 313 GLRAM-TGRIVGALASDKMPWVSLHFGTTHICGTLIDQWVLTAAHCFVTRKYL 371
Db 153 GRRLPVRIYGGQSSSLGRMPWVSLRYDGTLLCGSLGSDWVLTAAHCFVTRKYL 211
Oy 372 EGMKYVAG---TSNHLQLEPA--SIABIIINSY-----TDEEDYDIALMLSLPL 419
Db 212 SRWRFAFAVARTS-----PFAVQLGVQAVIYHGYLFPDPPTIDENGDIALVHLSSSL 266
Oy 420 TLSAHHHACLPMHGQTSLSMTCWITGFKTRERDDTSPFLARVQVNLIDFKKCNLY 479
Db 267 PLTEIIOVCLPAAQALVDGKVCVTGMMNT-QFYGOAAVVLQARVPILISNEVCNSPD 325
Oy 480 VYDSYLTFRMWCAGDLRGGRDSCQDSSGAPLVCEQ---NNRWYLAGVTSNGTCCGQRNK 535
Db 326 FYGNQIKRKMFCAGPRBGGIDACQDSSGHHFVCEBRIGTSRMRLCGIVSMGTGALARK 385
Oy 536 PGVYTKVTEVLPWYKMS 555
Db 386 PGVYTKVIDFERWIFQAIKT 405

RESULT 3

A56318
enteropeptidase (BC 3.4.21.9) precursor [validated] - human
N:Alternate names: enteropeptidase
C:Species: Homo sapiens (man)
C:Date: 19-May-1995 #sequence_revision 09-Aug-1996 #text_change 09-Jul-2004
C:Accession: A56318; B43090
R:Kitamoto, Y.; Velle, R.A.; Donis-Keller, H.; Sadler, J.E.
Biochemistry 34, 4562-4568, 1995
A:Title: cDNA sequence and chromosomal localization of human enteropeptidase, the proteolytic
A:Accession: A56318
A:Molecule type: mRNA
A:Residues: 1-1019 <KIT>

A:Cross-references: UNIPROT:P98073; UNIPARC:UPI000003FE65; GB:U09860; NID:9746412; PIDN:
R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A:Title: Enteropeptidase, the initiator of intestinal digestion, is a mosaic protease composed
A:Reference number: A43090; MUID:94329561; PMID:8052624
A:Accession: B43090

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 749-1019 <KIT>

A:Cross-references: UNIPARC:UPI0000172B0B; GB:U09860
C:Comment: The mechanism of association with the membrane of the intestinal brush border
ocated below) or with amino-terminal myristoylation of the heavy chain.

A:Gene: GDB:PRSS7

A:Cross-references: GDB:384083; OMIM:226200

A:Map position: 21q21-21q21

C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involve
ducts.

C:Function: cleaves activation peptide from trypsinogen to produce active trypsin

A:Description: intestinal digestive hydrolase cascade

A:Pathway: intestinal digestive hydrolase cascade

C:Superfamily: enteropeptidase; Ctr/Cts repeat homology; LDL receptor ligand-binding rep

C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen

F:1-784/Product: enteropeptidase heavy chain #status predicted <HCH>

F:12-38/Domain: transmembrane #status predicted <TMN>

F:184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:342-504/Domain: MAM homology <MAM>

F:526-631/Domain: Ctr/Cts repeat homology <CTR>

F:643-677/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:768-783/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC

F:785-1019/Product: enteropeptidase light chain #status predicted <LCH>

F:116-147/Domain: trypsin homology <TRY>

F:172-896,910-926,910-977,941-956,967-995/Dissulfide bonds: #status predicted

F:825,876,971/Active site: His, Asp, Ser #status predicted

Query Match 19.2%; Score 575; DB 1; Length 1019;

Best Local Similarity 33.3%; Pred. No. 5,56-24;

Matches 133; Conservative 58; Mismatches 156; Indels 52; Gaps 12;

Oy 187 HTGI-----RYKQRESCKRAVCDGVVDCYKSDDELQCFRPMDSKILIKYSGSSH 239
Db 636 HLGIPBCKADHFCQKNGECVPLVNLCDGHLHCBDSDEADCVRF-----FNGTNN 686
Oy 240 Q-----WPCSSNMNDSYEKTCQQLGFPSAHRTTEVAHRPANSFSLRYNS 288
Db 687 NGLVRFRIQSIWTTACAKEMTTQISNDVQQLLGSGSNSKPIFSTD---CGFVKLNT 743
Oy 289 TIOESL--HRSCEPSQRYISLQCSH--CGLRA---MTGRIVGALASDKMPWVSLH 339
Db 744 APDGLHILTPSQQLQSLIRLQCNHNSCGKLLAAODITPPIVGGSNKKEGAMPVVLVY 803
Oy 340 FGTTHICGTLIDQWVLTAAHCFVTRKYLEGMKYVAGTSLNHLQLEPA-----I 392
Db 804 YGRLLCGASLVSSDWSAHCYV--GRNLEPSKMTILG---DHMSNLTSPQTPRLI 859
Oy 393 AEIINSYVTEDEEDYDIALMLSLPLTSAHHPACLPMHGQTSLSMTCWITGFKTR 452
Db 860 DEIVINPHYNRKRRKNDIAMHLEFKVNTYTDYIPICLPENQVFPFRKNSISAGW--TV 918
Oy 453 ETDKTSPLAEVQVNLIDFKKNDYLVYDSYLTFRMWCAGDLRGGRDSCQDSSGAPLV 512
Db 919 VYGGTANILQEAADVPLSNRCCQNP--EYNITENNICAGYBEGGIDSCQDSSGAPLMC 977
Oy 513 EKNRWYLAGVTSNGTCCGQRNKRGVYTKVTEVLPWYKMS 551
Db 978 QENRWFLAGVTSNGYCALPFRGVARVARSFTETIOS 1016

RESULT 4

K0MSPL
plasma kallikrein (BC 3.4.21.34) precursor - mouse

C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Oct-2004
C:Accession: A36557
R:Seidman, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaudien, G.; Brachepapa, L.; Rochmon
DNA Cell Biol. 9, 737-746, 1990
A:Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and comparisc
A:Reference number: A36557; MUID:91090844; PMID:2264928
A:Accession: A36557
A:Molecule type: mRNA
A:Residues: 1-638 <SEI>
A:Cross-references: UNIPROT:P26262; UNIPARC:UPI00000277BC; GB:M58588; NID:9200358; PIDN:
A:Note: part of this sequence, including the amino ends of both the heavy and light chain
C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 11
C:Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-390/Product: plasma kallikrein heavy chain #status experimental <HCH>
F:20-109/Domain: apple repeat <AP1>
F:110-139/Domain: apple repeat <AP2>
F:200-289/Domain: apple repeat <AP3>
F:291-380/Domain: apple repeat <AP4>
F:391-638/Product: plasma kallikrein light chain #status experimental <LCH>
F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322
F:127,215,308,396,494/Binding site: carbonydrate (Asn) (covalent) #status predicted
F:434,483/Active site: His, Asp, Ser #status predicted
Query Match 19.1%; Score 574; DB 1; Length 638;
Best Local Similarity 33.7%; Pred. No. 3,9e-24;
Matches 131; Conservative 61; Mismatches 95; Indels 102; Gaps 15;
QY 196 RESCPKAVRCDGVV-----DCKLKSDLEG--VRFDWKSILKIYSSGHQWLPIC 246
DB 319 QETCK-TIRCOFPTYSLLPDCK---EEGCKSLRLSDSPRTIYV----- 363
QY 247 SNNMDSYEKTCQQLGFESAHRTTEVAHRPANSFSLRNSSTIOESLHSECSQRX 306
DB 364 MQSGSGYSLRLCKLV--DSPDCTTKI----- 387
QY 307 LQCSHGLRAMTGRIVGALASDSKMPQVSLH---FGTHICGTLIDAQWVLTAAICF 363
DB 388 -----NARIVGNTASLSGEMPMQVSLQVLTQVTHLCGSSIIGRQWVLTAAICF 436
QY 364 FVIREKYLEG-----WKVYAGTSLNHLQPE--AASIAEIIINSNYTDEEDVDIALM 414
DB 437 -----DGLPYDWMRIYSGIILSLSEITKETPESRIKEIIHQEYKVSSEGNVDIALIK 488
QY 415 LSKPLTISAHHPACLPMMHGQTSLSNETCWTIGFGRKRETDKTSPLREVOVNLIDPK 474
DB 489 LQPLANTTEPQKPLCLPSKADNTIYTNQWVGMWYKTE-QGETNIILOKATIPLPVPEE 547
QY 475 C-----NDVLVYDSYLTTPMMACAGDLRGGRDSCQSGGPGVCEONRWYLVAGTSMGTGC 530
DB 548 CQKTRDVTI-----NKQWTCAGYKEGGTDACKGDSGGLVCKHSGRQWLVGITSWEGSC 602
QY 531 GQNRKPGVYTKVTEVLPWYISKME-SEVR 558
DB 603 GRKQPGVYTKVSEYMDWILEKTQSSDVR 631
RESULT 5
KOHUP
Plasma kallikrein (EC 3.4.21.34) precursor - human
N:Alternate names: kininogenin; plasma prekallikrein
C:Species: Homo sapiens (man)
C:Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Oct-2004
C:Accession: A00921; A37939
R:Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
Biochemistry 25, 2410-2417, 1986
A:Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four p
A:Reference number: A00921; MUID:86243359; PMID:3521732

A:Accession: A00921
A:Molecule type: mRNA
A:Residues: 1-638 <CHU>
A:Cross-references: UNIPROT:P03952; UNIPARC:UPI000000DBAE; GB:M13143; NID:9190262; PIDN:
R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2050-2056, 1991
A:Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of 1
A:Reference number: A37939; MUID:91152016; PMID:1998666
A:Accession: A37939
A:Molecule type: Protein
A:Residues: 20-27;40-46,'X',48,'H',50,'X',52-70,'H',75-76,'X',78-80;103-113;131-140;141-1
260-283,'X',285;287-291,'X',293-295;314-317,'X',319-320;321-324,'X',329-333;334-339,'X',
525;338-551;562,'X',564-567;573,'X',575-576;578-583,'X',585;592-604 <MCX>
A:Cross-references: UNIPARC:UPI0000172B42; UNIPARC:UPI0000172B43; UNIPARC:UPI0000
10000172B52; UNIPARC:UPI0000172B53; UNIPARC:UPI0000172B54; UNIPARC:UPI0000172B55; UNIPARC
10000172B52; UNIPARC:UPI0000172B53; UNIPARC:UPI0000172B54; UNIPARC:UPI0000172B55; UNIPARC
C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 11
C:Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal re
inogen and may also play a role in the renin-angiotensin system by converting prorenin ir
C:Genetics:
A:Gene: GDB:KLM3
A:Cross-references: GDB:127575; OMIM:229000
A:Map position: 4q35-4q35
C:Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; inflan
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-638/Product: plasma kallikrein #status predicted <MAT>
F:20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>
F:20-109/Domain: apple repeat <AP1>
F:110-139/Domain: apple repeat <AP2>
F:200-289/Domain: apple repeat <AP3>
F:291-380/Domain: apple repeat <AP4>
F:391-638/Domain: plasma kallikrein light chain #status predicted <LCH>
F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383-
F:127,308,396,453,494/Binding site: carbonydrate (Asn) (covalent) #status experimental
F:318-347,340-345/Disulfide bonds: #status predicted
F:390-391/Cleavage site: Arg-11e (coagulation factor XIIa) #status predicted
F:434,483,578/Active site: His, Asp, Ser #status predicted
Query Match 19.1%; Score 572.5; DB 1; Length 638;
Best Local Similarity 34.5%; Pred. No. 4.7e-24;
Matches 132; Conservative 53; Mismatches 101; Indels 97; Gaps 13;
QY 196 RESCPKAVRCDGVV-----DCKLKSDLEG--VRFDWKSILKIYSSGHQWLPIC 245
DB 319 QETCK-MIRCOFPTYSLLPEDC--KEEKCKFLRLSDGSPTRIAYGTCSSGYSRLIC 375
QY 246 SNNMDSYEKTCQQLGFESAHRTTEVAHRPANSFSLRNSSTIOESLHSECSQRX 305
DB 376 NTGDNSTVCTTKI----- 387
QY 306 SLOCSHGLRAMTGRIVGALASDSKMPQVSLHFGTT---HICGTLIDAQWVLTAAICF 362
DB 388 -----SRIYVGTNNSGEMPMQVSLQVLTQVTHLCGSSIIGRQWVLTAAICF 435
QY 363 FVIREKYLEG-----WKVYAGTSLNHLQPE--AASIAEIIINSNYTDEEDVDIALM 413
DB 436 F-----DGLPYDWMRIYSGIILSLSDITKDPFSGIKETIHHQYKVSSEGNVDIALIK 487
QY 414 RLSKPLTISAHHPACLPMMHGQTSLSNETCWTIGFGRKRETDKTSPLREVOVNLIDPK 473
DB 488 KLOAPLANTTEPQKPLCLPSKADNTIYTNQWVGMWYKTEGE-IONTLQKVNIPLVNNE 546
QY 474 KNDYLVYDSY-LTPMMACAGDLRGGRDSCQSGGPGVCEONRWYLVAGTSMGTGCGQ 532
DB 547 ECQKR--YQDKITQRMWCAHYKGGKADCKGDSGGLVCHNGMRLVGTITSWEGCAR 604
QY 533 RNKQGVYTKVTEVLPWYISKMS 555
DB 605 RQPGVYTKVAEYMDWILEKTQS 627

```

membrane-bound arginine-specific serine proteinase precursor - rat
JC7731
C:Species: Rattus norvegicus (Norway rat)
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: UC7731; JC7775
C/Title: Characterization of a membrane-bound arginine-specific serine protease from rat
A:Reference number: UC7731; PMID:21421307; PMID:11530019
A/Accession: UC7731
A/Molecule type: mRNA
A/Residues: 1-855 <KIS>
A/Cross-references: UNIPROT:Q9JUI7; UNIPARC:UPI000008BAC0; DDBJ:AB049189
A/Experimental source: strain Male, 7-week-old
R:Satom, S.; Yamasaki, Y.; Tsuzuki, S.; Hitomi, Y.; Iwanga, T.; Fushiki, T.
Biochem. Biophys. Res. Commun. 287, 995-1002, 2001
A:Title: A role for membrane-type serine protease (MT-SP1) in intestinal epithelial turn
A:Reference number: JC7775; PMID:11573963
A/Contents: Small intestine
A/Accession: JC7775
A/Molecule type: mRNA
A/Residues: 1-855 <SAT>
A/Cross-references: UNIPARC:UPI000008BAC0; DDBJ:AB037898
C/Comment: This enzyme, an epithelial-derived, type II integral membrane serine protease
of specific proteins or peptides on the brushborder membranes. It also participates in
lial migration and/or cell loss.
C/Genetics:
A:Gene: mt-ep1
A:Map position: basolateral cell surface
A:Superfamily: membrane-bound arginine-specific serine proteinase
C/Keywords: protein digestion

Query Match      19.0%  Score 568.5; DB 2; Length 855;
Best local similarity 34.5%  Pred. No. 1e-23; Indels 81; Gaps 17;
Matches 141; Conservative 54; Mismatches 133;

QY 199 CPKHAIVCDGVDDCKLKSDELQCVFDMDSILKTYSSSHQWL-----P---ICGS- 247
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 464 CIRKDLRCDDGMADCPDYSDERHC-----RCNATQPFCKKQFCRPLFVDCSV 511

QY 248 -NMNDSYSEKTC-----OQLGFSAHRTTEVAHDPANSPSILRYN 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 512 NDCGGSGDEBEGSCSPAGSPKCSNGKCLPQSQCKNGKDCGGSDBASCDNNVAVSCTKYT 571

QY 288 STIGESL-----HRSCEPQORISLQCS-----HCLRAMT--GRVYGALASDSKMP 333
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 572 YRCQGLCLNKGNPCCDGRK---DSSGSDSEKNCDCGLRSFTYQARVVGATNADGEMP 627

QY 334 MGVSLH-FGTTTICGGLTIDAQWVLTAAHCF-----FVTRKVLKGMKVYAGTSLHQLP 387
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 628 MGVSLHAGGGLTLCASLISPDWLSVAHCFQDETIFKSDHTM--WTFPLGL--LDSK 683

QY 388 EAS-----IAETIINSYTDDEDDYIALMRSLKPLTSLAAHHPACLPHMGQTFSLNE 441
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 684 RSAGGVQEHKRIITHTSPSFDFTFDYDIALLELEKPAEYSTVAVPICLPDNTHTVPAGK 743

QY 442 TCMITPCKTERTDCKTSPFLREVOVNLIDFKKCDYLVVSYLTTPRMWCGDLRGSDS 501
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 744 AIWYTGWGHTRK-GGTGALLIOLKGRIRVINTOTTEILL--PQQTTPRMWCVGFLSGVDS 800

QY 502 CCGDSGGL-VCQNNRWYLAGVTSMTGCGQRNRPVGYTTKTEVLPWI 549
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 801 CCGDSGGLSVEVKGRITFGAGVVSWSBGCAQRNRKGYTTIIPVRDVI 849

```

```

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Oct-2004
C/Accession: A39180; A33320; S06851; I53041; S06852
R:Beaubien, G.; Rosinski-Chapin, I.; Mattei, M.G.; Mblakay, M.; Chretien, M.; Seidah, N.G.
Biochemistry 30, 1628-1635, 1991
A:Title: Gene structure and chromosomal localization of plasma kallikrein.
A/Reference number: A39180; PMID:91129236; PMID:1993180
A/Accession: A39180
A/Molecule type: DNA
A/Residues: 1-638 <BEA>
A/Cross-references: UNIPROT:P14272; UNIPARC:UPI00004ECC9; GB:J05315
A/Note: The authors translated the codon GAG for residue 81 as Gln
R:Seidah, N.G.; Lidenheim, R.; Mblakay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazure
DNA 8, 563-574, 1989
A:Title: The cDNA structure of rat plasma kallikrein.
A/Reference number: A33320; PMID:90091743; PMID:2598771
A/Accession: A33320
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-638 <SET>
A/Cross-references: UNIPARC:UPI00004ECC9; GB:M30282; NID:G205010; PIDN:AAA41463.1; PID:G
R:Paquin, J.; Benjannet, S.; Sawyer, N.; Lazure, C.; Chretien, M.; Seidah, N.G.
Biochim. Biophys. Acta 999, 103-110, 1989
A:Title: Rat plasma kallikrein: purification, NH(2)-terminal sequencing and development c
A/Reference number: S06851; PMID:90089457; PMID:2597701
A/Accession: S06851
A/Molecule type: protein
A/Residues: 20-45;391-413 <PAQ>
A/Cross-references: UNIPARC:UPI0000172B57; UNIPARC:UPI0000172B58
R:Seidah, N.G.; Lidenheim, R.; Mblakay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, R.; Lazure
DNA Cell Biol. 8, 563-574, 1989
A:Title: The cDNA structure of rat plasma kallikrein.
A/Reference number: I53041
A/Accession: I53041
A/Status: translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-638 <RSS>
A/Cross-references: UNIPARC:UPI000004ECC9; GB:M58590; NID:G206721; PIDN:AAA42069.1; PID:G
C/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 145
are linked by one or more disulfide bonds.
C/Genetics:
A:Gene: PK
C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; inflen
P:1-19/Domains: signal sequence; #statue predicted <SIG>
P:20-390/Product: plasma kallikrein heavy chain #statue experimental <MAT1>
P:20-109/Domains: apple repeat <AP1>
P:110-159/Domains: apple repeat <AP2>
P:200-289/Domains: apple repeat <AP3>
P:291-380/Domains: apple repeat <AP4>
F:391-638/Product: plasma kallikrein light chain #statue experimental <MAT2>
F:391-621/Domains: trypsin homology <TRY>
F:121-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,297-375,318-347,322-
F:121,215,308,453,459,494/Binding site: carbohydrate (asn) (covalent) #statue predicted
F:396/Binding site: carbohydrate (asn) (covalent) #statue experimental
F:433,483,578/Active site: His, Asp, Ser #statue predicted

Query Match      18.9%  Score 565.5; DB 1; Length 638;
Best local similarity 33.5%  Pred. No. 1.1e-23; Indels 107; Gaps 15;
Matches 130; Conservative 56; Mismatches 95;

QY 196 RESCPKHAIVCDGVV-----DCKLKSDELQCVFDMDSILKTI---YSGSSHQWL 243
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 319 QETCTK-TIRQFPYTSLLPDDCAE-----GCKSLRLSTGSPTRITYBAQSS----- 368

QY 244 ICSSNMNDSYSEKTCQQLGFSAHRTTEVAHDPANSPSILRYNSTIOESLHRECEPQR 303
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 -----GYSLRLCKV--ESSDCTTKI----- 387

QY 304 YISQSHGCLRAMTGRIVGALASDSKMPQVSLH---FGTTICGGLTIDAQWVLTAA 360
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 388 -----NARIVGSTNSLSGBMPQVSLQYKLVSQNMCGGSIIGROMILTAA 433

```

Db 1076 VCFSKVLGPGVYNSVSYFVGMI 1097

```

RESULT 10
A43090
enteropeptidase (EC 3.4.21.9) precursor [validated] - bovine
N/Alternate names: enterokinase
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: A43090; A48874; A61436
R/Kitamoto, Y.; Yuan, X.; Wu, O.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A/Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease compo
A/Reference number: A43090; MUID:94329561; PMID:8052624
A/Accession: A43090
A/Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1035 <KIT>
A/Cross-references: UNIPROT:P98072, UNIPARC:UPI000004BBB5, GB:U09859, NID:9746410, PIDN:
A/Experimental source: small intestine
R/Lavallee, E.R.; Rehemtulla, A.; Racine, L.A.; DiBasilio, E.A.; Ferenz, C.; Grant, K.L.;
J. Biol. Chem. 268, 23311-23317, 1993
A/Title: Cloning and functional expression of a cDNA encoding the catalytic subunit of k
A/Reference number: A48874; MUID:94043122; PMID:8226855
A/Accession: A48874
A/Molecule type: mRNA
A/Residues: 801-1035 <LAV>
A/Cross-references: UNIPARC:UPI00001133D, GB:L19663; NID:9416131; PIDN:AAA16035.1; PID:
A/Note: parts of this sequence, including the amino end of the mature protein, were conf
R/Light, A.; Janke, H.
J. Protein Chem. 10, 475-480, 1991
A/Title: The amino-terminal sequence of the catalytic subunit of bovine enterokinase.
A/Reference number: A61436; MUID:92189715; PMID:1799406
A/Accession: A61436
A/Molecule type: protein
A/Residues: 801-807, 'Y', 809-827 <LIG>
A/Cross-references: UNIPARC:UPI00001468A0
C/Comment: The mechanism of association with the membrane of the intestinal brush border
embrane attachment using a signal-anchor sequence.
C/Comment: Conversion from membrane-bound to soluble forms may involve further processin
C/Comment: mature enteropeptidase is variously reported to contain two (heavy and light)
1/ide linked
C/Function:
A/Description: cleaves propeptide from trypsinogen to produce active trypsin
A/Pathway: intestinal digestive hydrolase cascade
C/Superfamily: enteropeptidase; Clr/Cls repeat homology; LDL receptor ligand-binding reg
C/Keywords: glycoprotein; hydrolyase; intestine; serine proteinase; transmembrane protein
F/52-38/Domain: transmembrane #status predicted <TM>
F/52-117/Product: enteropeptidase mini chain #status predicted <MC>
F/118-800/Product: enteropeptidase heavy chain #status predicted <HC>
F/119-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F/159-520/Domain: MAM homology <MAM>
F/542-647/Domain: Clr/Cls repeat homology <CLR>
F/659-693/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F/694-799/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
F/801-1035/Product: enteropeptidase light chain #status predicted <LC>
F/801-1030/Domain: trypsin homology <TRY>
F/116,147,170,194,233,263,264,494,456,486,519,550,646,698,722,741,762,864,903,965/Bindin
F/788-912,826-862,926-993,957-972,983-1011/Diethylol bonds: #status predicted
F/841,892,987/Active site: His, Asp, Ser #status predicted

Query Match 18.3%; Score 548.5; DB 1; Length 1035;
Best Local Similarity 34.1%; Pred. No. 1.5e-22;
Matches 128; Conservative 54; Mismatches 144; Indels 49; Gaps 12;

QY 206 CDGVVDCUKLKDDELGCVRPDMKSLIKTYSGSSHQ-----WLPICSNMNDYS 254
DB 678 CDGPHCKDGSBEACVR-----LFGTTSSGLVQFRIGIMVACENMTTOIS 728
QY 255 EKTICCOLGFESAHRTTEVAHRDFANSFS-----ILRYNSTIOESL---HRSECPQARYIS 306
DB 729 DDVCCOLLGLGIGTGNSSVP-----TFETGGGPYVNLMTAPRGSLILTPSQCLEDSLIL 780
QY 307 LQCSH--CGLRAMT-----GRIVGALASDSKMPQVSLHFTTHICGGLTLDIAQNVLTAA 360

```

```

DB 781 LQCVKSCGKLVTOEVSPIKIVGSDSREGAMPVVALYFDQVCGASVSRDWLSAA 840
QY 361 HCFVTRKVLGVKKVYAG---TSNLQLP--EASINAEIINSVYTBEDDYDALMPLS 416
DB 841 HCIV-GRWMEPSKKKAVGLGMLMNSLTSPOLETRLIDIVINPYNKRKNNDIAMHLE 899
QY 417 KPLTSAIHHPACLPMHQGTFLNETCWITGFKTRTETDPTSPFLREOVNLIIDFKCN 476
DB 900 MKVNTVDYIGPICPEBNQVPPPRIGSIAGWG-ALTYQSTAVLVGADVPLLSNEKCO 958
QY 477 DLYVYDSYLTFRPMACAGDLRGGRSCGDSGGLVCEONNRYLAGVTSMGTGQQRKP 536
DB 959 QQMP-EVYITENMYCAGYEAGVDSGDSGGLPMCCENRMVLGAVTSFGYQCALPNRP 1017
QY 537 GVYTKVTEVLPWYS 551
DB 1018 GVIARVPRFTETWIOS 1032

```

```

RESULT 11
JC5759
brain-specific serine proteinase (EC 3.4.21.-) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 24-Jan-1998 #sequence revision 13-Mar-1998 #text_change 31-Dec-2004
C/Accession: JC5759
R/Yamamura, Y.; Yamashiro, K.; Tsunooka, N.; Nakazato, H.; Teujimura, A.; Yamaguchi, N.
Biochem. Biophys. Res. Commun. 239, 386-392, 1997
A/Title: Molecular cloning of a novel brain-specific serine protease with a kringlike-like
A/Reference number: JC5759; MUID:9608848; PMID:9344839
A/Accession: JC5759
A/Molecule type: mRNA
A/Residues: 1-761 <YAM>
A/Cross-references: UNIPARC:UPI0000175C77, DDBJ:D89871
A/Experimental source: brain
C/Superfamily: plasma hyaluronan-binding protein; scavenger receptor cysteine-rich domain
C/Keywords: glycoprotein; hydrolase; serine proteinase
F/85-157/Domain: kringlike-like #status predicted <KR>
F/163-266/Domain: scavenger receptor cysteine-rich domain homology <SRC>
F/166-266,273-372,386-486/Domain: scavenger receptor cysteine-rich #status predicted <SRC
F/513-516/Domain: furin binding #status predicted <FR>
F/517-755/Domain: trypsin homology <TRY>
F/93,521,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/562,612,711/Active site: His, Asp, Ser #status predicted

Query Match 17.9%; Score 536; DB 2; Length 761;
Best Local Similarity 24.5%; Pred. No. 5.4e-22;
Matches 189; Conservative 72; Mismatches 241; Indels 270; Gaps 29;

QY 17 AGASPAQASPAQTP---PGRASPAQASPAQTP-----PGRASPAQASPAQTP 65
DB 19 ARADPVSRSLPHRHPPSPRSQAHVLPSSRRPRTPRFLPLRIAPAQRFVLTGHTP 78
QY 66 ---PGRASPG-----RASPAQASPARASALSLSSSSGR 98
DB 79 PTIRRCAGSGMGWATNLGVPCILHWDVPPFLRLSRSPASVAELRGPHNCRBPDSGR 138
QY 99 -----SSASASATTSPTRYLVLRATPVGAV- 125
DB 139 PWCFTRNAQGVDMGYCDGCGPALPVIRLVGNSGHEG-----RVELTHAQMGWITC 191
QY 126 -----PIRSSPARSAPATRASTRSPGSLP 150
DB 192 DDQMDNADADVTCRQLGSLGIAKAMQAHFGESGPIILDLVRC-----TGNSMSIECCP 246
QY 151 KFTTRP---GOKPLPGCVLLIALVSLIILQPMWGHTGIRYKGRBS-CPKHAVRC 206
DB 247 KSSWGEHNGCHKEDAGVSCVPLTDG-VIRLAGKSTHEGRLEVYVYKQGMGTVCDDGWTM 305
QY 207 DGAVDCUKLKDDELGC-----CVAFDW----- 226
DB 306 NTYVACRLGKFKYKQSSVNHFPDSNRPIMLDDVSCSGKEVSFIQCSHRQGRHDCSRE 365

```

```

Oy 227 DKSL-----LKTVSGSSH-----OMLPICSSNMNDSEKTCQ 259
Db 366 DVLGLTCTPDSGHRISPEFPIRLVDGEEKKGRVEFVNGWGITCDGWTDKAAVTCR 425
Oy 260 QLGESAHRTTEVAHRDPANSFILRYNST-----IOESLRSECPQRYISL 307
Db 426 QLGKXGPARPRTMAY--FGEKGPIHMDNVKCTGNEKALADCVKDDIRHNCRHSEDAV 483
Oy 308 QC-----SHCGLRAM--TGRIVGALLADSKPMQVSLHFGTT-- 344
Db 484 ICDVLEKKASSGNKEMLSGGGLRLHRRQRRIIGVNNNSRGAMPQASILRLSAHSDG 543
Oy 345 --ICGGTIDQWVLTAAHCFVTRREKYLEGKQVYAGTSLNHL-----QLPEA 389
Db 544 RLTCGATLLSCWVLTAAHCF-----KRIQNNRSYANRVGDYHTLVPEEFQOE 592
Oy 390 ASIAEIIINSYVTEBEDYDIALMLSLKP-----LTLSAHHPACLP--HGQFSLNET 442
Db 593 IGVQOIVIHRYRPRRSYDIALVRLQPGRCARLSHVLPACLPLRERPRQKASN-- 650
Oy 443 CWTFEGKTRTBDKTSFPLREVOYNLIDFKKNDYLVYSYLPFRMCAGDLRGR--D 500
Db 651 CHITGWTGTRAYSRT--LQAAVPLLPKRFCKER--YKGLFTGRMLCAGNLQEDNRVD 705
Oy 501 SCQGDSSGGLVCEQ--NNRWYLAGVTSKGTGCGQRKPGVYTKTEVLPWYS 551
Db 706 SCQGDSSGGLVCEQ--NNRWYLAGVTSKGTGCGQRKPGVYTKTEVLPWYS 757

```

RESULT 12

```

KFNHU
coagulation factor Xla (EC 3.4.21.27) precursor (validated) - human
N/Alternate names: antihemophilic factor C; plasma thromboplastin antecedent
C/Species: Homo sapiens (man)
C/Date: 13-Aug-1986 #sequence revision 26-May-1994 #text_change 05-Oct-2004
A/Accession: A27431; A00920; A37940
R/Asekai, R.; Davie, E.W.; Chung, D.W.
Biochemistry 26, 7221-7228, 1987
A/Title: Organization of the gene for human factor XI.
A/Reference number: A27431; MUID:88107663; PMID:2827746
A/Accession: A27431
A/Molecule type: DNA
A/Residues: 1-625 <ASA>
A/Cross-references: UNIPROT:P03951; UNIPARC:UPI000000DB87; GB:M18295
A/Note: the sequence shown follows the authors' translation
R/Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
Biochemistry 25, 2417-2424, 1986
A/Title: Amino acid sequence of human factor XI, a blood coagulation factor with four te
A/Reference number: A00920; MUID:86243360; PMID:3636155
A/Accession: A00920
A/Molecule type: mRNA
A/Residues: 1-625 <FTU>
A/Cross-references: UNIPARC:UPI000000DB87; GB:M1142; NID:9182832; PIDN:AAA52487.1; PID:
Biochemistry 30, 2056-2060, 1991
R/McMullen, B.A.; Fujikawa, K.; Davie, E.W.
A/Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
A/Reference number: A37940; MUID:91152017; PMID:1998667
A/Accession: A37940
A/Molecule type: protein
A/Residues: 28-33335-49, 'X', 51-55, 'X', 58-63, 70-75, 'X', 77-79, 107-109, 'X', 111-112, 132-139;
1280-282, 'X', 284, 285-297, 313-316, 'X', 318-319, 320-326, 'X', 328-330, 'X', 347-349, 373, 'X', 375;
A/Cross-references: UNIPARC:UPI0000172824; UNIPARC:UPI0000172825; UNIPARC:UPI0000172826;
B/B; UNIPARC:UPI000017282C; UNIPARC:UPI000017282B; UNIPARC:UPI000017282B; UNIPARC:UPI000
10000172834; UNIPARC:UPI0000172835; UNIPARC:UPI0000172836; UNIPARC:UPI0000172837; UNIPAR
C/Comment: The proenzyme consists of two identical chains linked by one or more disulfid
he active site, and a heavy chain, which associates with high molecular weight (HMW) kin
C/Genetics:
A/Gene: GDB:F11
A/Cross-references: GDB:119891; OMIM:264900
A/Map position: 4q35-q35
A/Intron: 19/1; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 52
C/Function:
A/Description: catalyzes the proteolytic activation of coagulation factor IX

```

```

A/Pathway: blood coagulation intrinsic pathway
C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C/Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro
F/1-18/Domain: signal sequence, duplication; glycoprotein; hemophilia C; homodimer; hydro
F/19-387/Product: coagulation factor Xla heavy chain #status experimental <HCH>
F/119-108/Domain: apple repeat <AP1>
F/109-198/Domain: apple repeat <AP2>
F/199-288/Domain: apple repeat <AP3>
F/290-379/Domain: apple repeat <AP4>
F/388-625/Product: coagulation factor Xla light chain #status experimental <LCH>
F/388-625/Domain: trypsin homology <TRY>
F/20-103, 514-581, 571-599/Disulfide bonds: #status predicted
F/29/Disulfide bonds: interchain #status experimental
F/46-76, 50-56, 110-193, 136-165, 140-146, 200-283, 226-255, 230-236, 291-374, 317-346, 321-327, 380
F/90, 126, 353, 450/Binding site: carbohydrate (asn) (covalent) #status predicted
F/339/Disulfide bonds: interchain #status predicted
F/387-388/Cleavage site: Arg-116 (coagulation factor XIa) #status experimental
F/431, 480, 575/Active site: His, Asp, Ser #status predicted
F/491/Binding site: carbohydrate (asn) (covalent) #status experimental

```

```

Query Match 17.7%; Score 532; DB 1; Length 625;
Best Local Similarity 33.2%; Pred. No. 7, 2e-22;
Matches 131; Conservative 62; Mismatches 127; Indels 74; Gaps 17;

```

```

Oy 212 CLKSDLEGG--CYRPMDKSLKTVSGSSHQWLP--C-SSNMND-----SYSE 255
Db 255 CLKTSBGLPSTIKKSKALSGSLSCRSIFVCHSSYHDTDFGLBELDIAAKSH 314
Oy 256 KTCQOLGFESAHRTTEVAHRDPANSFILRYNSTIOESLRSECPQRYISLQCSHGLR 315
Db 315 EACQKL-----CTNAVRCQF--FTYTPAASCNBEG--KGKC---YLKLSNGSPTK 358
Oy 316 AMTG-----RIVGALLADSKPMQVSLHFGTT-----HIC 346
Db 359 ILHRSGGISGTYLRLCKMNECTTKIKPRIVYAGTSLNHLQPEAS--IAEIIINSYV 416
Oy 347 GGTLLDAQWVLTAAHCF--VTRREKYLEGKQVYAGTSLNHLQPEAS--IAEIIINSYV 402
Db 417 GGTLLDAQWVLTAAHCF--VTRREKYLEGKQVYAGTSLNHLQPEAS--IAEIIINSYV 402
Oy 403 DEBDYDIALMLSLKPPLTSLAHHPACLP--HGQFSLNETGTFGKTRTBDKTSFPL 462
Db 474 MAESGYDIALMLSLKPPLTSLAHHPACLP--HGQFSLNETGTFGKTRTBDKTSFPL 462
Oy 463 REVQVNLIDFKKNDYLVYSY--LTPRMCMGDLRGHSCQGSGLVCEBQNNRWYLA 521
Db 522 GVTSMGTCGCGQRKPGVYTKTEVLPWYS 555
Oy 522 GVTSMGTCGCGQRKPGVYTKTEVLPWYS 555
Db 591 GVTSMGTCGCGQRKPGVYTKTEVLPWYS 555

```

RESULT 13

```

PIBO
plasma (EC 3.4.21.7) precursor - bovine
N/Alternate names: plasminogen
C/Species: Bos primigenius taurus (cattle)
C/Date: 30-Sep-1987 #sequence revision 28-Apr-1995 #text_change 09-Jul-2004
A/Accession: S45046; A25835; I45961; S03736
R/Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A/Description: Cloning and characterization of the bovine plasminogen cDNA.
A/Reference number: S45046
A/Accession: S45046
A/Molecule type: mRNA
A/Residues: 1-812 <BER>
A/Cross-references: UNIPROT:P06668; UNIPARC:UPI0000043BBB; EMBL:X79402; NID:9494962; PID:
A/Experimental source: liver
A/Note: it is uncertain whether Met-1 or Met-8 is the initiator
R/Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Roselet, S.J.; Kampfer, U.; Rickl
Eur. J. Biochem. 149, 267-278, 1985
A/Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasmi

```

A:Reference number: A25835; MUID:85203906; PMID:3846532

A:Accession: A25835

A:Molecule type: protein

A:Residues: 27-334; D' ,336-S1S,'H','517-554,'L','556-812 <SCH>

A:Cross-references: UNIPARC:UIP0000172B95

F:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.

Biochemistry 23, 4243-4250, 1984

A>Title: Characterization of a complementary deoxyribonucleic acid coding for human and

A:Reference number: IAS961; MUID:85023311; PMID:6148961

A:Accession: IAS961

A>Status: translated from GB/EMBL/DDBT

A:Molecule type: mRNA

A:Residues: 706-743;'R','745-812 <MAL>

A:Cross-references: UNIPARC:UIP000016C365; GB:K02935; NID:g163551; PIDN:AAA30714.1; PID:

R.Brundholm, R.A.; Lerch, P.G., Schaller, U.; Rickti, E.B.; Lergier, W.; Manneberg, M,

Eur. J. Biochem. 114, 465-470, 1981

A>Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,

A:Reference number: S03735; MUID:81212097; PMID:7238497

A:Accession: S03736

A:Molecule type: protein

A:Residues: 27-83 <BRU>

A:Cross-references: UNIPARC:UIP0000172B96

C:Function:

A>Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety
of the walls of the graatian follicle; also activates the urokinase-type plasminogen acti-

A:Pathway: fibrinolysis

G:Superfamily: plasmin; kringe homology; plasminogen-related protein precursor homology

C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolases; kidney; kringe; plasma,

P:1-26/Domains: signal sequence #status predicted <SIG>

P:8-103/Domains: plasminogen-related protein precursor homology <PLPH>

P:27-812/Product: plasminogen #status experimental <PRO>

P:127-103/Domains: activation peptide #status experimental <APT>

P:104-583,584-812/Product: plasmn #status experimental <MAT>

P:110-188/Domains: plasmin chain A #status experimental <ACH>

F:1192-269/Domains: kringe homology <KR1>

F:1282-359/Domains: kringe homology <KR2>

F:1384-461/Domains: kringe homology <KR3>

F:1485-554/Domains: kringe homology <KR5>

F:1584-812/Domains: plasmin chain B #status experimental <BCB>

F:1584-805/Domains: trypsin homology <TRY>

F:156-80, 60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,334-
461 bonds: #status predicted

F:315/Binding site: carbohydrate (asn) (covalent) #status experimental

F:365/Binding site: carbohydrate (ser) (covalent) #status experimental

F:624,667,762/Active site: His, Asp, Ser #status predicted

```

DB      766  LVCFEKDKYILGGVTSWGLGCAKRNKPGVYVVRSPVYPMTEBETM 809

RESULT 14
A57014
proctasin (EC 3.4.21.-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: A57014; A54866
R:Yu, J.X.; Chao, L.; Chao, J
J. Biol. Chem. 270, 13483-13489, 1995
A:Title: Molecular cloning, tissue-specific expression, and cellular localization of human
A:Reference number: A57014; MUID:95286644; PMID:7768952
A:Accession: A57014
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-343 <RES>
A:Cross-references: UNIPROT:Q16651; UNIPARC:UPI0000046DCB; GB:L41351; NID:9862304; PIDN:;
A:Experimental source: prostate
A>Note: parts of this sequence were determined by protein sequencing
R:Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 269, 18843-18848, 1994
A:Title: Proctasin is a novel human serine proteinase from seminal fluid. Purification,
A:Reference number: A54866; MUID:94308140; PMID:8034638
A:Accession: A54866
A:Molecule type: protein
A:Residues: 45-64 <YUA>
A:Cross-references: UNIPARC:UPI0000172AFB
C:Genetics:
A:Gene: GDB:PRS68
A:Cross-references: GDB:676446; OMIM:600823
A:Map position: 16p11.2-16p11.2
C:Superfamily: trypsin, trypsin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-44/Domain: product: proctasin #status predicted <ANT>
F:33-44/Domain: proctasin light chain #status predicted <CHL>
F:45-343/Domain: proctasin heavy chain #status predicted <CHH>
F:45-281/Domain: trypsin homology <TRY>
F:323-341/Domain: transmembrane #status predicted <TMN1>
F:37-154/70-86,168-244,201-223,234-262/Disulfide bonds: #status predicted
F:85,134,238/Active site: His, Asp, Ser #status predicted
F:159/Binding site: carbohydrate (Aen) (covalent) #status experimental

Query Match      17.0%; Score 510.5; DB 1; Length 343;
Best Local Similarity 40.9%; Pred. No. 5.7e-21;
Matches 103; Conservative 42; Mismatches 94; Indels 13; Gaps 5;

QY      312  CGLRMTGTVIGGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFPTREKVL 371
DB      37   CGV-APQARITGGSSAVAGQMPQVSIITYGVAHCGSLVSEQVLSAAHCFPSEHNR-- 93
QY      372  EGMKYVAGTSNLTQLEPAASIA--EIIINSNYTDEEDYDIALMLRSLPTLSAHIHRA 428
DB      94   EAYEVKLGAHQDVSSEDAKVSFLKDIIHPSPVLSQESSQGDIALQLSRPTFSRYIRPI 153
QY      439  CLPHNGQTSLNLETGTCITGFGKRETRDDKTSPP-IRSVQNLIDFKCNQDLYLVDS---- 483
DB      154  CLPAAANASPPNGLHCTVTGWMGHVAPSVSLITPRLPQLQLEVPILSRFTCNCLYNIDAKPEE 213
QY      484  --YLTPRMWACAGDILNGRSDSCQDSSGAPLVCEQNNRWYLAGVTSWGTGCGQRNKPQVYTK 541
DB      214  PHFVGEDWVACAVYVBEGSKACQGDSSGPLSCPVEGLMYLGLGISWMDACGARNRPGVYTL 273
QY      542  VTEVLPMYYSKM 553
DB      274  ASSYASWIOSKV 285

RESULT 15
UC4171
trypase (EC 3.4.21.59) precursor - rat
N:Alternate names: mast cell trypase

```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:05:05 ; Search time 301 Seconds
(without alignments)
1727.106 Million cell updates/sec

Title: US-10-806-370-12
Perfect score: 2999
Sequence: 1 MERUSHGNSAPARTPSAGAS.....TEVLPMIYSKMESEVFRPKS 562

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 7.2:*
1: uniprot_sprot:*
2: uniprot_tramb1.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------|
| 1 | 2956 | 98.6 | 581 | 1 | TMPSD_HUMAN |
| 2 | 2510.5 | 83.7 | 543 | 1 | TMPSD_MOUSE |
| 3 | 891 | 29.7 | 359 | 2 | Q4SPGQ_TETNG |
| 4 | 777.5 | 25.9 | 490 | 1 | TMPS2_MOUSE |
| 5 | 771.5 | 25.7 | 490 | 2 | Q3UKK3_MOUSE |
| 6 | 770.5 | 25.7 | 490 | 2 | Q7TN04_MOUSE |
| 7 | 769.5 | 25.7 | 490 | 2 | Q6P7D7_RAT |
| 8 | 763.5 | 25.5 | 453 | 2 | Q812A6_MOUSE |
| 9 | 763.5 | 25.5 | 475 | 2 | Q3T206_MOUSE |
| 10 | 763.5 | 25.5 | 490 | 2 | Q920K3_RAT |
| 11 | 762.5 | 25.4 | 453 | 1 | TMPS3_MOUSE |
| 12 | 762.5 | 25.4 | 453 | 2 | Q2M1G4_MOUSE |
| 13 | 761.5 | 25.4 | 484 | 2 | Q311U4_MACMU |
| 14 | 759.5 | 25.3 | 486 | 2 | Q5PRA6_BRARE |
| 15 | 755 | 25.2 | 484 | 2 | Q311V4_PANPA |
| 16 | 754 | 25.1 | 484 | 2 | Q311V3_PANTR |
| 17 | 751 | 25.0 | 484 | 2 | Q311V5_PANTR |
| 18 | 744 | 24.8 | 484 | 2 | Q311U8_HYLSY |
| 19 | 740.5 | 24.7 | 492 | 2 | Q6GTLK_HUMAN |
| 20 | 738.5 | 24.6 | 492 | 1 | TMPS2_HUMAN |
| 21 | 738.5 | 24.6 | 492 | 2 | Q96T73_HUMAN |
| 22 | 736.5 | 24.6 | 538 | 2 | Q5USC7_HUMAN |
| 23 | 731 | 24.4 | 454 | 1 | TMPS3_HUMAN |
| 24 | 713.5 | 23.8 | 722 | 2 | Q6NUE5_XENLA |
| 25 | 712.5 | 23.8 | 767 | 2 | Q9DGR2_XENLA |
| 26 | 697.5 | 23.3 | 437 | 1 | TMPS4_HUMAN |
| 27 | 696 | 23.2 | 445 | 2 | Q8CJ17_RAT |
| 28 | 689.5 | 23.0 | 435 | 2 | Q5RDX7_PONPY |
| 29 | 684.5 | 22.8 | 435 | 1 | TMPS4_MOUSE |
| 30 | 684 | 22.8 | 388 | 2 | Q4RRR7_TETNG |
| 31 | 668.5 | 22.3 | 371 | 2 | Q8CJ16_RAT |

| | | | | | | |
|----|-------|------|-----|---|--------------|---------------------|
| 32 | 664 | 22.1 | 455 | 1 | TMPS5_MOUSE | Q9ER04_mus musculus |
| 33 | 659 | 22.0 | 455 | 2 | Q8CDR0_MOUSE | Q8CDR0_mus musculus |
| 34 | 644.5 | 21.5 | 457 | 1 | TMPS5_HUMAN | Q9H3A3_homo sapien |
| 35 | 616.5 | 20.6 | 445 | 2 | Q3U0U6_MOUSE | Q3U0U6_mus musculus |
| 36 | 611.5 | 20.4 | 417 | 1 | HEPS_HUMAN | P05981_homo sapien |
| 37 | 609.5 | 20.3 | 417 | 2 | Q5R5E8_PONPY | Q5R5E8_pongo pygma |
| 38 | 608.5 | 20.3 | 436 | 1 | HEPS_MOUSE | Q35453_mus musculus |
| 39 | 600 | 20.0 | 326 | 2 | Q7Z280_BRARE | Q7Z280_brachydantio |
| 40 | 592.5 | 19.8 | 416 | 1 | HEPS_RAT | Q05511_rattus norv |
| 41 | 580.5 | 19.4 | 730 | 2 | Q4RH70_TETNG | Q4RH70_tetradon n |
| 42 | 578.5 | 19.3 | 572 | 1 | TMPS7_MOUSE | Q8B1K6_mus musculus |
| 43 | 578 | 19.3 | 799 | 2 | TMPS6_MOUSE | Q9D10_mus musculus |
| 44 | 578 | 19.3 | 799 | 2 | Q6PFA4_MOUSE | Q6PFA4_mus musculus |
| 45 | 578 | 19.3 | 811 | 2 | Q3KN88_MOUSE | Q3KN88_mus musculus |

ALIGNMENTS

RESULT 1
TMPSD_HUMAN STANDARD, PRT, 581 AA.
ID TMPSD_HUMAN Q99YE2; Q86YM4; Q960Y8; Q99YE1;
AC Q99YE2; Q86YM4; Q960Y8; Q99YE1;
DT 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT 15-MAR-2005, sequence version 2.
DT 07-MAR-2006, entry version 29.
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
DE protease) (Membrane-type mosaic serine protease).
GN Name=TMPS13; Synonyms=MSP, TMPS11;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RX MEDLINE=21167393; PubMed=11267681; DOI=10.1016/S0167-4781(01)00184-1;
RC TISSUE=Lung;
RA Kim D.R., Sharmis S., Inoue M., Kido H.,
RT "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung."
RL Biochim. Biophys. Acta 1518:204-209(2001).
[2]
RA NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2).
RP Pak T.J., Park W.J.,
RN "Homo sapiens transmembrane protease, serine 6 (TMPS6) mRNA";
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
[3]
RA NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4).
RC PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi W., Kanda K., Yokoi T., Furuya T., Kikkawa B., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niimiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kinata M., Matsumoto M., Hirao K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano Y.,
RA Kanehori K., Takahashi F., Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togaya S., Komai F., Hara R., Takeuchi K., Ariizawa N., Imose N.,
RA Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,
RA Yodaishino K., Yuzuki H., Oshima A., Sasaki N., Aotaka S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shintzu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiiwara Y.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsunuma K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Matanabe M., Konatsu T.,
 RA Mizushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs".
 RL Nat. Genet. 36:40-45(2004).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RG Human chromosome 11 international sequencing consortium,
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
 CC protein (Potential).
 CC -1 ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=MSP1, Large form, Membrane-type;
 CC IsoId=Q9BYE2-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9BYE2-2; Sequence=VSP_013103, VSP_013104;
 CC Note=No experimental confirmation available;
 CC Name=3; Synonyms=MSPS, Small form;
 CC IsoId=Q9BYE2-3; Sequence=VSP_013099, VSP_013102;
 CC Name=4;
 CC IsoId=Q9BYE2-4; Sequence=VSP_013100, VSP_013101;
 CC Note=No experimental confirmation available;
 CC -1 TISSUE SPECIFICITY: Isoform 1 and isoform 3 are predominantly
 CC expressed in lung, placenta, pancreas, and prostate. Isoform 3 is
 CC weekly expressed in testis and peripheral blood lymphocytes.
 CC -1 POLYMORPHISM: The repeat A-S-P-A-(GLQR) is polymorphic and the
 CC number of copies varies between 12 to 14.
 CC -1 SIMILARITY: Belongs to the peptidase S1 family.
 CC -1 SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -1 SIMILARITY: Contains 1 SRCK domain.
 CC -1 CAUTION: Ref.2 has referred to this protein as TMPRSS6.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License

CC -----
 CC EMBL; AB048796; BAB39741.1; -; mRNA.
 CC EMBL; AB048797; BAB39742.1; -; mRNA.
 CC EMBL; AY190317; AAC38062.1; -; mRNA.
 CC EMBL; AK027798; BAB5376.1; -; mRNA.
 CC EMBL; AP002962; -; NOT_ANNOTATED_CDS; Genomic_DNA.
 CC HSRP; P00760; IEXX.
 CC MEROPS; S01.087; -;
 CC EMBL; ENSG00000137747; Homo sapiens.
 CC HGNC; HGNC:29808; TMPRSS13.
 CC GO; GO:0016021; C:integral to membrane; NAS.
 CC GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
 CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
 CC InterPro; IPR002172; LDL_rcpt_A.
 CC InterPro; IPR001254; Peptidase_S1_S6.
 CC InterPro; IPR001314; Peptidase_S1.
 CC InterPro; IPR001190; Srcr_rcpt_
 CC Pfam; PF00057; Ildl_rcpt_a; 1.
 CC Pfam; PF00089; Trypsin_1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS01209; LDLRA_1; FALSE_NEG.
 CC PROSITE; PS00068; LDLRA_2; FALSE_NEG.
 CC PROSITE; PS00420; SRCR_1; FALSE_NEG.
 CC PROSITE; PS00287; SRCR_2; 1.
 CC PROSITE; PS0240; TRYP_SIN_DOM; 1.
 CC PROSITE; PS00134; TRYP_SIN_HIS; 1.
 CC PROSITE; PS00135; TRYP_SIN_SER; 1.
 CC Alternative splicing; Glycoprotein; Hydrolase; Membrane; Polymorphism;
 KW Protease; Repeat; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 581 Transmembrane protease, serine 13.
 FT TOPO_DOM 1 160 Cytoplasmic (Potential).
 FT

FT TRANSMEM 161 181 Signal-anchor for type II membrane
 FT TOPO_DOM 182 581 protein (Potential).
 FT REPEAT 9 13 Extracellular (Potential).
 FT REPEAT 14 18 2-1 (approximate).
 FT REPEAT 19 23 1-2.
 FT REPEAT 24 28 1-3.
 FT REPEAT 29 33 2-2.
 FT REPEAT 34 38 1-4.
 FT REPEAT 39 43 1-5.
 FT REPEAT 44 48 1-6.
 FT REPEAT 49 53 2-3.
 FT REPEAT 54 58 1-7.
 FT REPEAT 59 63 1-8.
 FT REPEAT 64 68 2-4.
 FT REPEAT 69 73 1-9 (approximate).
 FT REPEAT 74 78 1-10.
 FT REPEAT 79 83 1-11.
 FT REPEAT 84 88 1-12.
 FT DOMAIN 199 320 SRCR.
 FT DOMAIN 199 221 LDL-receptor class A.
 FT REGION 321 554 Peptidase S1.
 FT REGION 9 88 12 x 5 AA repeats of A-S-P-A-(GLQR).
 FT COMPBIAS 14 68 4 x 5 AA repeats of T-P-P-G-R.
 FT ACT_SITE 361 361 Ala-rich.
 FT ACT_SITE 409 409 Charge relay system (By similarity).
 FT ACT_SITE 506 506 Charge relay system (By similarity).
 FT CARBOHYD 287 287 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 400 400 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 440 440 N-linked (GlcNAc...) (Potential).
 FT DISULFID 245 309 By similarity.
 FT DISULFID 258 312 By similarity.
 FT DISULFID 346 362 By similarity.
 FT DISULFID 443 512 By similarity.
 FT DISULFID 475 491 By similarity.
 FT DISULFID 502 530 By similarity.
 FT VARSPLIC 146 181 GTSLPKPTWRGQQLPLIGCVLLILVALVSLILF -> V
 FT (in isoform 3).
 FT VARSPLIC 423 486 /FtId=VSP_013099.
 FT VARSPLIC 487 581 AHHPACLPMTGQFTSLMETCWINGFGKTRTDDTKSPFLR
 FT EGVQVLIDFKKNDLVLDVSLT -> GEGICTPSPAPQP
 FT OHPLPSHLASVNSVSPGPKASAGQSKTLDPVMEHFCFI
 FT IRETEAGCL (in isoform 4).
 FT /FtId=VSP_013100.
 FT VARSPLIC 487 581 Missing (in isoform 4).
 FT /FtId=VSP_013101.
 FT VARSPLIC 555 581 VRSIQDTPAFRLGTSGGDPGAPRV -> SEVAFRKS
 FT (in isoform 3).
 FT /FtId=VSP_013102.
 FT VARSPLIC 555 588 VRSI -> SSAG (in isoform 2).
 FT /FtId=VSP_013103.
 FT VARSPLIC 559 581 Missing (in isoform 2).
 FT /FtId=VSP_013104.
 FT CONFLICT 187 187 H -> Y (in Ref. 2).
 FT CONFLICT 192 192 Y -> H (in Ref. 1; BAB39741).
 FT CONFLICT 201 201 K -> E (in Ref. 2).
 FT CONFLICT 206 206 C -> R (in Ref. 2).
 FT CONFLICT 259 259 C -> R (in Ref. 1).
 FT CONFLICT 298 298 E -> H (in Ref. 1).
 FT CONFLICT 298 298 H -> R (in Ref. 2).
 FT CONFLICT 496 496 H -> R (in Ref. 2).
 FT SEQUENCE 581 AA; 62680 MW; A4935CCCF31D29EE CRC64;
 SQ

Query Match 98.6%; Score 2956; DB 1; Length 581;
 Best Local Similarity 99.8%; Pred. NO. 3.3e-155;
 Matches 553; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MERDSHGASPARTSPASGASPAQASPAQASPAQASPAQASPAQAS 60
 Db 1 MERDSHGASPARTSPASGASPAQASPAQASPAQASPAQASPAQAS 60
 Qy 61 PAGTPPGASPGASPAQASPARASPALASLSRSSSGSSGRSASASVTTSPTRYVLYVAT 120

```

Db 61 PAGPFGASPGASPAQASPARASPALASLSRSSSSGSSSSSARSASVTSPRYLVAT 120
Qy 121 PVGAVPITSSPARASPARATRESPTSLPKPTWEEGQKOLPLGCVLLIALVSLITL 180
Db 121 PVGAVPITSSPARASPARATRESPTSLPKPTWEEGQKOLPLGCVLLIALVSLITL 180
Qy 181 FQFMGHTGIRKRESCRESCPKIAVACDGVDCCKLSDELGCYRPFMDLSLTKYSGSSHQ 240
Db 181 FQFMGHTGIRKRESCRESCPKIAVACDGVDCCKLSDELGCYRPFMDLSLTKYSGSSHQ 240
Qy 241 WLPICSSNMWDSYSBKTCQOLGFESEAHRTTEVAHDPANSPSILKYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMWDSYSBKTCQOLGFESEAHRTTEVAHDPANSPSILKYNSTIOESLHRSCEP 300
Qy 301 SQRYSISLQCSHCGLRAMGRIVGALADSKRPMQVSLHFGTTHCGSTLLDAQVVLTA 360
Db 301 SQRYSISLQCSHCGLRAMGRIVGALADSKRPMQVSLHFGTTHCGSTLLDAQVVLTA 360
Qy 361 HCFPVTREKVLKGMKVAGTSLNHLQPEASIAEIIINSNTDEEDYDIALMRLSKPLT 420
Db 361 HCFPVTREKVLKGMKVAGTSLNHLQPEASIAEIIINSNTDEEDYDIALMRLSKPLT 420
Qy 421 LSAHIHPACLPMHGTFSINETCWTGFKTRETDDKTSPLREYQVNLIDFKKNDYLV 480
Db 421 LSAHIHPACLPMHGTFSINETCWTGFKTRETDDKTSPLREYQVNLIDFKKNDYLV 480
Qy 481 YDSYLTPTPMACAGLGRGDSGQSGPLVCEONNRVYLAGVTSWGCGRNRPQYVT 540
Db 481 YDSYLTPTPMACAGLGRGDSGQSGPLVCEONNRVYLAGVTSWGCGRNRPQYVT 540
Qy 541 KYTEVLPWYSKME 554
Db 541 KYTEVLPWYSKME 554

```

RESULT 2

```

TMPSPD MOUSE STANDARD; PRT: 543 AA.
ID TMPSPD MOUSE STANDARD; PRT: 543 AA.
AC Q5U405; Q8CFB0; Q9IV08;
DT 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT 15-MAR-2005, sequence version 2.
DT 07-MAR-2006, entry version 16.
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
DE protease) (Membrane-type mosaic serine protease).
GN Mus musculus; Synonyms=Msp;
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
NM NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=B6/EGFP; and FVB/N;
RC TISSUE=Mammary tumor, and Trophoblast stem cell;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Alekshun S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Murina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,
RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

```

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
CC protein (Potential).
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
CC -1- SIMILARITY: Contains 1 SRCR domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL: BC010843; AAH10843.1; -; mRNA.
DR EMBL: BC042878; AAH42878.1; -; mRNA.
DR EMBL: BC085323; AAH85323.1; -; mRNA.
DR HSSP: P00760; IEZX.
DR MEROPS: S01.087; -.
DR MGI: MGI:2682935; Tmpres13.
DR InterPro: IPR002172; LDL_rcpt_A.
DR InterPro: IPR001254; Peptidase_S1_S6.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR001190; Srcr_rcpt.
DR Pfam: PF00057; Ldl_recept_a; 1.
DR Pfam: PF00530; SRCR; 1.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00202; SR; 1.
DR SMART: SM00202; Tryp_Spc; 1.
DR PROSITE: PS01209; LDLRA_1; FALSE_NEG.
DR PROSITE: PS00668; LDLRA_2; FALSE_NEG.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS0287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR GlycoProtex: Hydrolase; Membrane; Protease; Repeat; Serine protease;
KW Signal-anchor; Transmembrane.
KW CHAIN 1 543
FT 1 TOPO_DOM 1 143
FT 2 TRANSMEM 144 164
FT 3 REPEAT 165 543
FT 4 REPEAT 14 17
FT 5 REPEAT 18 22
FT 6 REPEAT 23 27
FT 7 REPEAT 28 31
FT 8 REPEAT 32 36
FT 9 REPEAT 37 40
FT 10 REPEAT 41 45
FT 11 REPEAT 46 49
FT 12 REPEAT 50 54
FT 13 REPEAT 55 59
FT 14 REPEAT 60 64
FT 15 REPEAT 65 69
FT 16 DOMAIN 180 202
FT 17 DOMAIN 199 301
FT 18 DOMAIN 302 535
FT 19 REGION 14 49
FT 20 REGION 18 69
FT 21 COMBPTS 11 72
FT 22 ACT_SITE 342 342
FT 23 ACT_SITE 390 390
FT 24 ACT_SITE 487 487
FT 25 CARBOHYD 231 231
FT 26 CARBOHYD 268 268
FT 27 CARBOHYD 381 381
FT 28 CARBOHYD 421 421
FT 29 DISULFD 226 290
FT 30 DISULFD 239 293
FT 31 DISULFD 327 343
FT 32 DISULFD 424 493

```


AC O9108:99JK4; O9QY82;
 DT 11-JAN-2001, integrated into UniProtKB/Swiss-Prot.
 DT 11-JAN-2001, sequence version 2.
 DT 07-MAR-2006, entry version 52.
 DE Transmembrane protease, serine 2 (EC 3.4.21.-) (Epitheliasin) (Plasmic
 DE transmembrane protein X) [Contains: Transmembrane protease, serine 2
 DE non-catalytic chain; Transmembrane protease, serine 2 catalytic
 DE chain].
 GN Name=TmpRSS2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], AND TISSUE SPECIFICITY.
 RC STRAIN=BALB/c;
 RX DOI=10.1002/1096-9896(2000)9999:9999::AID-PATH743>3.0.CO;2-T;
 RA Vavrala M.H., Potvrat R.S., Kellokumpu S., Kyllonen A.P., Vihko P.T.;
 RT "Expression of transmembrane serine protease TMPRSS2 in mouse and
 RT human tissues.";
 RL J. Pathol. 193:134-140(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RA Han J., Kim S.;
 RT "Putative transmembrane protease X.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RC STRAIN=BALB/c;
 RX PubMed=10683448; DOI=10.1016/S0014-5793(00)01196-0;
 RA Jacquinet E.J., Rao N.V., Rao G.V., Hoidal J.R.;
 RT "Cloning, genomic organization, chromosomal assignment and expression
 RT of a novel mosaic serine proteinase: epitheliasin.";
 RL FEBS Lett. 468:95-100(2000).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=129; TISSUE=Mammary gland;
 RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.W., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.F., Casavant T.L., Schetz T.E.,
 RA Stepien M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Morley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Bakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalske U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
 CC cleavage and secreted (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed mainly in prostate and kidney.
 CC -1- SIMILARITY: Belongs to the peptidase S1 family.
 CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL, AF199362; AAF97867.1; -, mRNA.
 CC EMBL, AF243500; AAF64186.1; -, mRNA.
 DR

[illegible]

Db 47 -----SPVQYAPRITTOASTSVIHTHPKSGCAPCTSKSKS 83
 Qy 145 POTSIPKFTWRREGOKLPILGCVLLIALVSLIILFPOMOCHTGRYKQRES--CPKH 202
 Db 84 LCLALALGT-----VLTGAVAVALILMRWDSNCSTSEMECGSSSTCISG 128
 Qy 203 AVRCDGVVDCKLKSDGLGCVRFDMKSLIKIYSGSSHOWLPICSSNMNDYSSEKTCOOLQ 262
 Db 129 SLMGCGVNAHCPNGEDENRCVRLYQGSFLLQVSSQRKAMYPICODMSSYRAACXDMG 188
 Qy 263 FESARITTE-VAHRPANSFSLIRYNS---TIQSLHRSBSPQRYISLQCSHCGLRAM 317
 Db 169 KKNYNSSGGIPDQSGATSFMELNVSSGNVDLYKLLYHSDSCSSRNVSLSRCTIEGVSSV 248
 Qy 318 --TGRIVGALASDCKMPQVSLHFGTTHICGGTILIDAMOVYLAACFVTEKVLG-- 373
 Db 249 KRQSTIVGGLANSRQDMQVSLHGVVCGSITTEPMYTAHCV---EELPLSGPR 304
 Qy 374 -WKVYAGT-----SNLHQLPEAASIAEIIINSNTYDEDDYIALMRSLKPLTLSAH 424
 Db 305 YVTAFAGILRQSLMFYSGRHQ-----VEKVISHPYDSOKTGNIDALMKLQTPLAFNDL 358
 Qy 425 IHPACLPNMGQTSINERTWTITGFKTRTDKTSRPLAEVQVNLIDPKKCNVDLYVYSY 484
 Db 359 VKPVLPPNGWMLDDDBQBCWISGWCATYE-KGKTSDVNAAMVPLIEBCKSKSYIYNL 417
 Qy 485 LTPRMWCAADDLRGSDSCGSDGSPVCGVGNRRWYLAQVTSNGTGGORNKPGVYTKYTE 544
 Db 418 ITPAMICAGFLQGSYDSCGSDGSGPLVTLKNGIMWLGDTSGSCAKALRGVIGNVTV 477
 Qy 545 VLPWYISKMS 555
 Db 478 FTDWYIQMRA 488
 RESULT 5
 Q3UKES MOUSE
 ID Q3UKES3 MOUSE PRELIMINARY; PRT; 490 AA.
 AC Q3UKES3
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2005, sequence version 1.
 DT 21-EB-2006, entry version 7.
 DE 14 days pregnant adult female placenta cDNA, RIKEN full-length
 DE enriched library, clone:153026G04 product:transmembrane protease,
 DE serine 2, full insert sequence.
 DE Name:Tmpres2;
 CN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carinici P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RL Methode Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carinici P., Kasukawa T., Katayama S., Gough J., Fitch M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Batiz M.B., Bremer S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Aldins V., Allen J.R.,
 RA Banerji I., Impombato A., Apweiler R., Attaliya R.N., Bailey T.L.,
 RA Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Flecher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojbbori T., Green R.E.,
 RA Guentrich S., Harbers M., Hayashi Y., Henrich T.K., Hirokawa N.,
 RA Hill D., Hummel L., Iacono M., Ikeo K., Iwano J., Ishikawa T.,
 RA Jaki M., Kampin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kito M., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Laveau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Maderazo M., Marchionni L.,
 RA Matsumura H., Matsuzawa S., Miki H., Mignone F., Miyake S., Moritz K.,
 RA Moravskiy T., Mulder N., Nakano N., Nakamura H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Noi F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Patasi G., Pezole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Roest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sezer L., Sheng Y.,
 RA Shibata Y., Shimada E., Shimada K., Silva D., Sinclair B.,
 RA Spelling S., Stupka E., Sugiyama K., Sultana R., Takenaka Y., Taki K.,
 RA Tamaki K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen B., Verardo R., Wei C.B., Yeig K.,
 RA Yamashita H., Zdobych B., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Matlick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imanura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Kono H., Nakano K., Nimomiya N.,
 RA Nihei T., Okada M., Plesey C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome";
 RL Science 309:1559-1563(2005).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RA RIKEN Genome Exploration Research Group, and Genome Science Group
 RG (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome";
 RL Science 309:1564-1566(2005).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi U., Bono H., Kondo S.,
 RA Nishikido I., Ose T., Saito R., Suzuki H., Yamashita I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Haegawa Y., Nogami A., Schonbach C., Gojbbori T.,
 RA Balderick L.M., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirini L.M., Kanapin A., Matsumura H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frezer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guentrich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petter G., Pezole G.,
 RA Petrovsky N., Pillai R., Pontius U.V., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Secou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yangisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carinici P., Hayatsu N.,
 RA Hirozawa-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume K., Imocani K., Ithil Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komno H., Aachi J., Fukuda S.,
RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kabukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Maturo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guentrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohlsunki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN (6)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN (7)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Komno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara B., Wachihi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN (8)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Komno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Hydrolase; Membrane; Protease; Serine protease; Transmembrane.
SQ SEQUENCE 490 AA; 5349 MW; 54651AF288E76EBA C6C64;
Query Match 25.7%; Score 771.5; DB 2; Length 490;
Best Local Similarity 31.5%; Pred. No. 6,6e-35;
Matches 174; Conservative 87; Mismatches 200; Indels 91; Gaps 13;
OY 27 AGTPGGRASPQASPAQASPAQTPGGRASPQASPAQTPGGRASPQASPAQASPA---R 83
DB 5 SGSPFG-IGPCENHGYSEHICPP-----RPAPANGYNLPQGYSPVQ 51
OY 84 ASPLASISRSSSSSSSSARSASVTSPTRYLVATPVGAVPVIRSSPARSAPATRYRE 143
DB 52 YAPRTTYQASTSVYTHHPKSSGALCTSSKSLCLALALGV----- 93
OY 144 SPGTSLPFTWRREGQQLPLIGCVLLLTALVLSLILFPQMGHTGRYKQRES--CPK 201
DB 94 -----LTGAAVAVLNRFPMDNSNSTSEMEGSGTGCS 127
OY 202 HAVRCDGVNDCKLSDELGVCFPMWKSLLKIYSGSSHQWLPICSSNNNDYSEKTCQQL 261
DB 128 SSLWCDDVAAHCPNGEDENRCVRLYGQSFLLQVYSSQRAKWPVPCDDWSESYGRAACKDM 187
OY 262 GPESHARTE-VAHKDFANSFSLRYNS---TIOESLHRECEPQRYISLQCHGCLR 316
DB 188 GYKNFVSSQGIPODSGATSMKLNVSNGVNLVKKLHSDSCSSRMVSLRCIEGVR 247
OY 317 M--TGRIVGALASDPKPMQVSLFPTTHCGTLLDAQWLVLAHCFVTRERVLG- 373
DB 248 VKRGRIVGLNABSGMPQVSLHVGAVHVGSGIITPEMIVTRAHCV---EPLSP 303
OY 374 --WKYVACT-----SNLQLEPAASIAIINSNTYDEEDYDALMRSLKPLTSA 423
DB 304 RYWTAFACILQSLMFFYSRQ-----VEKVISHPNYSKYKNDIALMKLQTLAFND 357
OY 424 HIHPACLEPMAGTBSLNETCWITGKTRERDDKTSPLREYQVNLIDPKKNDVLVYDS 483
DB 358 LVKPVCLPSPGMMLDLDECEWISGAGATYE-KGKTSVDVLAAMVLLIPSKNSRYLYN 416
OY 484 YLTPRMAGDLRGGRDSQGDSPGLVCEONRMYLVAGTYSWGTGCGORNRKPYTKT 543
DB 417 LITPMTCAGFLQGSVDSQDGGSPVLTAKNGIWMVLTGDSWSSGCAKALRPYGVNT 476
OY 544 EYLPWYSKMS 555
DB 477 VFTDWIKQMR 488
RESULT 6
OY 07TN04 MOUSE PRELIMINARY; PRT; 490 AA.
AC 07TN04-
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DE 21-FEB-2006, entry version 16.
DR Transmembrane protease, serine 2.
GN Name=Trypsin2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129; TISSUE=Mammary tumor, Brca1-/f1;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold B.A., Grove L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marnusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Staphleion M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollahy S.J.,
 RA Boeak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettlemen M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=129; TISSUE=Mammary tumor. Brcal-/fl;
 RA Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: BC054348; AAH54348.1; -; mRNA.
 DR HSSP: P00760; 1EZX.
 DR Ensemble: ENSMUSG00000000385; Mus musculus.
 DR MGI: MGI:1354381; Tmpres2
 DR GO: GO:0016021; C: integral to membrane; RCA.
 DR InterPro: IPR0012172; LDI_rcpt_A.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00057; Ldl_recept_a; 1.
 DR Pfam: PF00530; SRCR_1.
 DR Pfam: PF00089; Trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00192; LDLA_1.
 DR SMART: SM00202; SR_1.
 DR SMART: SM00202; TRYP_SPC_1.
 DR PROSITE: PS01209; LDLRA_1; 1.
 DR PROSITE: PS50068; LDLRA_2; 1.
 DR PROSITE: PS50287; SRCR_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SRR; 1.
 KW Hydrolyase; Membrane; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 490 AA; 53526 MW; 54650B028417655A CRC64;
 Query Match 25.7%; Score 770.5; DB 2; Length 490;
 Best Local Similarity 31.5%; Pred. No. 7.5e-35;
 Matches 174; Conservative 87; Mismatches 200; Indels 91; Gaps 13;
 Oy 27 AGTPGRASPAPQASPAQASPACTPPGRASPAPQASPACTPPGRASPAPQASPA---R 83
 Db 5 SGPSPG-IGPCVENGYSEHICP-----PPVAPNPNYNYPPQYSPVQ 51
 Oy 84 ASPALASIRSSSGSSSARSASAVTSPRYLVATVGVAVPISSPARASAPARATRE 143
 Db 52 YAPRTTQASISVHTHPRKSGALCTSKSKSLCALALGV----- 93
 Oy 144 SPGTSLPFRWRBQKQPLIGCVLLILALVSLILFQFOGHTGIRYKQRES--CPK 201
 Db 94 -----LQGAANVAATLWRPFMDNSNSTSEMGSSGTCIS 127
 Oy 202 HAVKCDGVVDCKLKSDLCVRFPMDKSLTKYSSSHQMLPICGSMNNDISSEKTCOL 261
 Db 128 SSLWCDDVAHCPNGEDENRCVRLYGQSPILQYSSQKAWYFVCCDDWSBSYGRACADM 187
 Oy 262 GFESAHRTTE-VAHRDFANSFELIKYNS---TIGESLHRSQPCSORIYSIQCSHCGR 316
 Db 188 GYKNNFYSSQGIPODSQATSPKLVNVSNGVNDLYKLYHSDSCSSRMVYSIRCEGVRS 247

Oy 317 M--TGRIVGALASDPSKMPQVSLHFGTTHICGTLIDAQWLTAAHCFVTRKYLEG- 373
 Db 248 VKRGRITVGINASGDMPMQVSLHVGQVHVGGSIIIPENIVTAHCV---EPLSSP 303
 Oy 374 --WKVYACT-----SNLHQLPEASIAEIIINSNTYDEEDYDIALMRISKPLTUSA 423
 Db 304 RWTAFATILRQSLMFYGRHQ-----VEKVISHPNYSKTKKNDIALMKLQTPLEAFND 357
 Oy 424 HIHPCLPMHQTFELNETCWITGPKTRERDDKTPPLREYQVLLIDPKKNDLVYDS 483
 Db 358 LKVPCLPFGMMLLDDECMISGWATYE-KGKTSVLAANAAPLIPSKNSKYITNN 416
 Oy 484 YLTPRMQAGLRRGGRDSCQSDSGPLVCEONNRYLAGVTSWGTGCGQARNKPGVYTYT 543
 Db 417 LITPAMICAGFLQGSVDS CQSDSGPLVTLNKGTWLLIGDTSWGSCKALRPGYGVNT 476
 Oy 544 EYLPWYISKMS 555
 Db 477 VFTDWIYQOMRA 488
 RESULT 7
 Q6P7D7_RAT PRELIMINARY; PRT; 490 AA.
 ID Q6P7D7;
 AC Q6P7D7;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 21-FEB-2006, entry version 18.
 DE Transmembrane protease, serine 2.
 GN Name=Tmpres2;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Eumetazoa; Gyltre; Rodentia; Sciurognathi;
 OC Mammalia; Eutheria; Eumetazoa; Gyltre; Rodentia; Sciurognathi;
 OC Muroidea; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel H.,
 RA Diatchenko L., Marnusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Staphleion M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollahy S.J.,
 RA Boeak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettlemen M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: BC061712; AAH61712.1; -; mRNA.
 DR HSSP: P00761; 1AKS
 DR Ensemble: ENSMUSG000000001976; Rattus norvegicus.
 DR GO: GO:0016021; C: integral to membrane; IEA.

OY 494 DLRGGRDSCQSDSGPLVCEQNNRWYLAGVTSWGTGCGQRNPGVYTKTEVLPMIYGM 553
 DB 388 YLKGVDSCQSDSGPLVCEQNNRWYLAGVTSWGTGCGQRNPGVYTKTEVLPMIYGM 447
 OY 554 ESEVR 558
 DB 448 ERDLK 452
 RESULT 9
 Q3T206_MOUSE PRELIMINARY, PRT: 475 AA.
 AC Q3T206;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2005, sequence version 1.
 DT 21-FEB-2006, entry version 7.
 DE Adult inner ear cDNA, RIKEN full-length enriched library,
 clone:R930041F06 product:transmembrane protease, serine 3, full insert
 sequence.
 GN Name=Trps3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 CC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Fritch M.C., Maeda N.,
 Oyama R., Ravasi T., Lenhard R., Wells R., Kodzius R., Shimokawa K.,
 Bajic M.B., Brenner S.B., Batalov S., Forrest A.R., Zavolan M.,
 Davis M.J., Wilm L.G., Altshuler D., Lenhard R., Altshuler D., Bailey T.L.,
 Ambesi-Impombato R., Apweiler R., Aurrelija R.N., Bailey T.L.,
 Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 Chiu K.P., Choudhury V., Chruschke D.R., Chruschke D.R.,
 Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
 Fletcher C.F., Fukushima T., Furuno M., Furukawa S., Gariboldi M.,
 Georgi-Hemming P., Gingeras T.R., Gojobori T., Green R.B.,
 Guignard S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
 Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 Jaki M., Kanapin A., Katoh M., Kawasawa Y., Keise J., Kitamura H.,
 Kitoan H., Kohler G., Kishino S.P., Kruger A., Kummerfeld S.K.,
 Kurochkin I.V., Lazareu L.F., Lazarevic D., Lipovich L., Liu J.,
 Liuti S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 Mutsaers J., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Okazaki Y., Ohtsuka V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 Roest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 Schombach C., Sekiguchi K., Sempile C.A., Sessa L., Sheng Y.,
 Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 Spelling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
 Tamada K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 Ueda H.R., van Nimwegen B., Vercato R., Wei C.L., Yang K.,
 Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide W., Bult C.,
 Gilmord S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 Wamboldt C., Matlock J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimura S.,
 Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
 Tagami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,

RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563 (2005).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi Y., Bono H., Kondo S.,
 Nishikawa Y., Ohtsuka V., Saito R., Suzuki H., Yamashita H.,
 Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,
 Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 Schmitt L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
 Gasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
 Gilmord S., Guignard S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 Konagaya A., Kurochkin I.V., Lee Y., Lenhard R., Lyons P.A.,
 Megiort D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
 Petrovsky N., Piliat R., Pontius J.U., Qi D., Ramchandran S.,
 Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 Verardo R., Wagner L., Wamboldt C., Wang Y., Watanabe Y., Wells C.,
 Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang L.,
 Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamashita H.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gasterland T., Giesi C., King B., Kochwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
 Schmitt L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Guignard S., Hill D., Hofmann M., Hume D.A., Kanai A., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Matzke R., Mombere P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schombach C., Seya T., Shibata Y., Storch K.F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschütter S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes."
RN Genome Res. 10:1617-1630(2000).
[17]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Inner ear;
RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/9r.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama K., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishie T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa K., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer."
RN Genome Res. 10:1757-1771(2000).
[18]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Inner ear;
RA Aikawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Horii F., Iida Y., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Kono H., Murata M., Nakamura M., Niimura N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shizaki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
CC EMBL; AK158203; BAB34404.1; -; mRNA.
CC MGI; MGI:2155445; Tmpres3.
DR GO: GO:0016021; C:integral to membrane; RCA.
DR GO: GO:0016020; C:membrane; RCA.
DR GO: GO:0004263; F:chymotrypsin activity; RCA.
DR GO: GO:0016787; F:hydrolyase activity; RCA.
DR GO: GO:0005044; F:scavenger receptor activity; RCA.
DR GO: GO:0004295; F:trypsin activity; RCA.
DR GO: GO:0006508; P:proteolysis; RCA.
DR InterPro; IPR002172; LDI_rcpt_A.
DR InterPro; IPR001254; Peptidase_S1_96.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_rcpt_Pfam; PF00057; Ldl_recept_a: 1.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLA_1.
DR SMART; SM00202; SR_1.
DR SMART; SM00202; Tryp_Spc_1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS01069; LDLRA_2; 1.
DR PROSITE; PS02877; SRCL_2; 1.
DR PROSITE; PS02040; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KM Hydrolyase; Membrane; Protease; Serine protease; Transmembrane.
SQ SEQUENCE 475 AA; 51886 MW; 0214D1B550EC917B CR664;
Query Match 25.5%; Score 763.5; DB 2; Length 475;
Best Local Similarity 38.6%; Pred. No. 1.8e-34;
Matches 164; Conservative 68; Mismatches 150; Indels 43; Gaps 13
QY 161 LPL-----ICGULLALAVSLILIPF---WGCHTIRKREGRSPKNAVRDQIV 210
DB 66 LPLKFFPIIVIGIALLIALAIGLHIFGCSKRYRHSFK-----CIEFLARCDGVS 118
QY 211 DCKLSDSLGCVRFPMDSILKITYSGSSHOWLPICSSNNNDYSSEKTCQOLF-----SSA 266
DB 119 DCKNADEYRCVRCVGGQRAALQVFLNAA--WRTMGSDPKSHYAKAKAQOLFSPVSVD 176
QY 267 HRTTEVAHRDPANFSILRYNSTIOE---SLH-----RSECPSPQRYISLQSCGGLRAM 317

```

Db      177 HLRVLDLEQFOGDF--VSINHLLSDDKVTALHHSVYRRECGTSGHVVTLKCSACGR-- 223
Qy      318 TG---RIVGALASDSSKPMQVSLHFGVTHICGGLIDAQVVLTAACFFVTRKVLG 314
Db      233 TGVSPRIVGGMSSSLTQMPQVSLQFGVHLGGSIIPFLIVTAACVVDLYHP--KSW 290
Qy      375 KYVAGTSNHLQPEASIALF-IITNSNTDEEDVDIALMLSLKPLSLAHIPALCPM 433
Db      291 TVQGVGLVSLMDSPVPSHLVEKIIYHSKYPKRLGNDIALMLSLPLFDETIQICLPNS 350
Qy      434 GQFFSLNEFCWLTGCKTRETDDKTSPLRREVQVNLIDFKKCNQVLYVDSVLFPRM 493
Db      351 EENFPDGLKCMWISGMKATEDGGD-ASPVLNHAAPVLISNKCINRDVYGGIISMLCAG 409
Qy      494 DLRGKSDSCQSGSGGPLYVCEQNNRWVYLAGVTSWGTGCGQRNKPQVYTKVTEVLP 553
Db      410 YLKGVDSCQSGSGGPLYVCEQERLMLKVLGATSFVIGCAEVNKKPGVYRITSF 466
Qy      554 ESEVR 558
Db      470 ERDLK 474

RESULT 10
Q920K3_RAT PRELIMINARY; PRT; 490 AA.
AC Q920K3;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 21-FEB-2006, entry version 24.
DE TPFRS52.
GN Name=TPFRs52; Synonyms=TPFRS52;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Tezuka S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
CC protein (By similarity).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
Cc
Cc EMBL; AB073350; BAB70683.1; -; mRNA.
Cc DR HSSP; P00760; 1EZK.
Cc DR MEROPS; S01.247; -.
Cc DR Ensembl; ENSRNOG0000001976; Rattus norvegicus.
Cc
Cc RGD; 620766; TPFRS52.
Cc DR GO; GO:0016020; C:membrane; IEA.
Cc DR GO; GO:0005044; F:scavenger receptor activity; IEA.
Cc DR GO; GO:0004552; F:serine-type endopeptidase activity; IEA.
Cc DR GO; GO:0006008; P:protein catabolism; IEA.
Cc DR InterPro; IPR002172; LDL_rcpt_A.
Cc DR InterPro; IPR001254; Peptidase_S1_S6.
Cc DR InterPro; IPR001314; Peptidase_S1A.
Cc DR InterPro; IPR001190; Src_rcpt.
Cc Pfam; PF00057; Icdl_recept_a; 1.
Cc Pfam; PF00089; Trypsin; 1.
Cc DR PRINTS; PRO0722; CHYMOTRYPSIN.
Cc DR SMART; SM00192; LDLa; 1.
Cc DR SMART; SM00202; SR; 1.
Cc DR SMART; SM00020; TRYP_Spc; 1.
Cc DR PROSITE; PS50068; LDLRA_2; 1.
Cc DR PROSITE; PS50287; SRCR_2; 1.
Cc DR PROSITE; PS50240; TRYPSIN_DOM; 1.
Cc DR PROSITE; PS00135; TRYPSIN_SER; 1.
Cc DR Hydrolase; Membrane, protease; Serine protease; Transmembrane.
Cc SSOURCE 490 AA; 53519 MW; 286691551CAC4094 CRC64;

```



```

OY 211 DCKLDELGVRPDMKSLIKIYSGSSHOWLPICSSNMNDSYSEKTCQQLGF-----BSA 266
DB 97 DCKNAEDRYRCVRSGQPAALQVFTAA--WRTWCSDMKSHYAKIACAQGFPEYVSSD 154
OY 267 HRTTEVAHRDPANSPSILRYNSTIOE---SLH-----RSECPQRYISLQCSHGLPAM 317
DB 155 HLRVDALSEQFGDPF--VSINHLSDDKYTALHSHVYMRBECTSGHVVTLKCSACGTR-- 210
OY 318 TG---RIYGALASDSKMPQVSLHFGTTHICGGLIDAOVYLTAAHCFVTRREKVLBGW 374
DB 211 TGYSPRIVGMMSSLTQMPQVSLQFGYHLCGSGSVITPLMIVTAACVYDLYHP--KSW 268
OY 375 KYVAGTSLHQLPEASIAE--IINSNTDEEDVDIALMRSLKPLTSAHHPACLPMH 433
DB 269 TVQVGLVSLMDSPPVSHLVEKIYHSKYPKRLGNDIALMLKSEPLTDEDTIOPICLNS 328
OY 434 GQTSINETCMTITGFKTRTDDKTSPPFLREVQVNLIDFKCNDLYVDSYLTPEPMCGAG 493
DB 329 EENFPDGLCMTISGKATEDEGD--ASPVLNHAAVPLISNKNCHRDVYGGIISPMCLAG 387
OY 494 DLRGGRSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGQRNKPQYVTVTEVLPWYISKM 553
DB 388 YLKGQVDSQGDGSGPLVCEQERRMLKVLGATSFGLGCAEVNKPQYVTVITISFLDWIHQL 447
OY 554 ESEVR 558
DB 448 ERDLK 452

RESULT 12
O2MIG4_MOUSE PRELIMINARY; PRT; 453 AA.
ID O2MIG4_MOUSE AC
AC O2MIG4;
DT 21-FEB-2006, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, sequence version 1.
DE Transmembrane protease, serine 3.
GN Name=TMPRSS3;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
CN [1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=PCR rescued clones;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bostak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC112375; AAI12376.1; -, mRNA.
DR KMW Protease; Transmembrane.
SQ SEQUENCE 453 AA; 49492 MW; 1ABCBF10AF6E1EF6 CRC64;

Query Match      25.4%; Score 762.5; DB 2; Length 453;
Best Local Similarity 38.6%; Pred. No. 1.9e-34;
Matches 164; Conservative 68; Mismatches 150; Indels 43; Gaps 13;

OY 161 LPL-----IGCVLLALAVLSLILPQF---WQSGHGRYKQRESCPKHAAVCDGVV 210
DB 44 LPLKFPPIVIGITALLIALIGLHFDGCGKTRCHSSFR-----CIELTRACDVS 96
OY 211 DCKLDELGVRPDMKSLIKIYSGSSHOWLPICSSNMNDSYSEKTCQQLGF-----BSA 266
DB 97 DCKNAEDRYRCVRSGQPAALQVFTAA--WRTWCSDMKSHYAKIACAQGFPEYVSSD 154
OY 267 HRTTEVAHRDPANSPSILRYNSTIOE---SLH-----RSECPQRYISLQCSHGLPAM 317
DB 155 HLRVDALSEQFGDPF--VSINHLSDDKYTALHSHVYMRBECTSGHVVTLKCSACGTR-- 210
OY 318 TG---RIYGALASDSKMPQVSLHFGTTHICGGLIDAOVYLTAAHCFVTRREKVLBGW 374
DB 211 TGYSPRIVGMMSSLTQMPQVSLQFGYHLCGSGSVITPLMIVTAACVYDLYHP--KSW 268
OY 375 KYVAGTSLHQLPEASIAE--IINSNTDEEDVDIALMRSLKPLTSAHHPACLPMH 433
DB 269 TVQVGLVSLMDSPPVSHLVEKIYHSKYPKRLGNDIALMLKSEPLTDEDTIOPICLNS 328
OY 434 GQTSINETCMTITGFKTRTDDKTSPPFLREVQVNLIDFKCNDLYVDSYLTPEPMCGAG 493
DB 329 EENFPDGLCMTISGKATEDEGD--ASPVLNHAAVPLISNKNCHRDVYGGIISPMCLAG 387
OY 494 DLRGGRSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGQRNKPQYVTVTEVLPWYISKM 553
DB 388 YLKGQVDSQGDGSGPLVCEQERRMLKVLGATSFGLGCAEVNKPQYVTVITISFLDWIHQL 447
OY 554 ESEVR 558
DB 448 ERDLK 452

RESULT 13
O31IU4_MACMU PRELIMINARY; PRT; 484 AA.
ID O31IU4_MACMU AC
AC O31IU4;
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Transmembrane protease serine 2 (Fragment).
GN Name=TMPRSS2;
OS Macaca mulatta (Rhesus macaque);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
CX NCBI_TaxID=9544;
CN [1]
RN NUCLEOTIDE SEQUENCE.
RP Clark N.L., Swanson W.J.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; DOI:50506; AA282295.1; -, Genomic DNA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0008233; F: peptidase activity; IEA.
DR GO; GO:0005044; F: scavenger receptor activity; IEA.
DR GO; GO:0004252; F: serine-type endopeptidase activity; IEA.

```

DR GO:0006508; P:proteolysis; IEA.
 DR InterPro: IPR002172; LDL_rcpt_A.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Srcr_rcpt.
 DR Pfam: PF00057; LDL_recept_a; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00192; LDLa; 1.
 DR SMART: SM00202; SR; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS01209; LDLRA_1; 1.
 DR PROSITE: PS00068; LDLRA_2; 1.
 DR PROSITE: PS50287; SRCR_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Protease; Transmembrane.
 KM NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 484 AA; 53082 MW; 09E3EBD8739792B0 CRC64;
 Query Match 25.4%; Score 761.5; DB 2; Length 484;
 Best Local Similarity 33.3%; Pred. No. 2,3e-34;
 Matches 185; Conservative 81; Mismatches 188; Indels 101; Gaps 19;
 QY 28 GTPPGR-----SPAQASPAQASPAQTPPG--RASPQASPAQTPPGRASPGRASPA 77
 DB 1 GSPPVGVYENHGQVNPYPAPQPTVA---PNVEVHPAPQVYPPV-----QYTPR 50
 QY 78 QASPARASPALASLSRSSGSSGSSASVTTSPRYVLVTRATPVGAVPINSSPARASPA 137
 DB 51 VLTTH-SNPAAVCRQPKSPSG---TVCTSKTKKALCVMTLGAV----- 89
 QY 138 TRATRESGTSLPKRTWRGQKQLPLIGCVLLTLVLSIILPQFMQ-----GHTGIR 191
 DB 90 -----LVDA-----ALAAGL-----WKFMGSKSDSGLS 114
 QY 192 YKQGRSCPKHAIVCDGVVDCKLKSDELGCYVFPWDKSLIKIYSGSSHQMLPTICSSNMND 251
 DB 115 C-DSSGCTISSNMCDGVSHCPNGEDENRCVRLVGPNNILQVYSQSRSMHPVCDDNNE 173
 QY 252 SYSEKTCQQLGPESAHRTTE-VAHKDFANSFSLRY---NSTIGSLHRSF-CPSQRYTS 306
 DB 174 NYAABAACRMQGVKNSFVSSOGIVDSGATSPFKLNTSAGNDIYKKLYHSDACSSKAYVS 233
 QY 307 LQCHGGR---AMTGRIVGALASDSKPMQVSLHFTTHICGTLIDAGMVLTAACF 363
 DB 234 LRCTACGRNLSRROSRIVGQONALLGAMPQVSLHVNHCSSITTPETIVTAACV 293
 QY 364 FVTEKVLG---WKVYAGTSLNHLQ--PFAASIAEIIINSNYTDEDDYDIALMRLSKP 418
 DB 294 ----EKLPSNPMQMTAFVGTILRSSMFLPKGRVAKVISHNYVDSKTNQNDIALMKLTP 349
 QY 419 LTLSAHHPACLPNMGQTFSLNETCWITGFKTRETDTKTSPPLEVOVNLIDFKCNDY 478
 DB 350 LTFEIVVAVPLCLPNMGMLLEPQHMWISGMGATQF--KCKTSDVLAANAAPPILPEPRCNK 408
 QY 479 LVYDSYLTPRMWCAGLDGRDSCGDSGGLVCEONNRWTLAIGTSGTCCGQGRKGV 538
 DB 409 YVYGLITPAHICAGFLQGVDSCGDSGGLVTLKNDVMMILGDTSGSGCAANRRGV 468
 QY 539 YTKVTEVLPWYIKM 553
 DB 469 YGNVTVFTDWIYRQM 483
 RESULT 14
 QSPRA6 BRARB PRELIMINARY; PRT; 486 AA.
 AC QSPRA6
 DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
 DT 04-JAN-2005, sequence version 1.

DT 21-FEB-2006, entry version 10.
 DE Zgc:101791.
 GN ORFNames=zgc:101791;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=olfactory epithelium;
 RX MEDLINE=22389257; PubMed=12479932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Alechuth S.F., Zeeberg B., Bluetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditchenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitoki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paley J., Heltun B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalske U., Smallue D.R.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=olfactory epithelium;
 RX Director MGC Project; to the EMBL/GenBank/DBJ databases.
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
 CC protein (By similarity).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL: BC086738; AA086738.1; -; mRNA.
 DR Ensemble: ENSDARG0000045889; Danio rerio.
 DR ZFIN: ZDB-GENE-041212-48; Zgc:101791.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0005044; F:scavenger receptor activity; IEA.
 DR GO: GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO: GO:0006508; P:proteolysis; IEA.
 DR InterPro: IPR002172; LDL_rcpt_A.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Srcr_rcpt.
 DR Pfam: PF00057; LDL_recept_a; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00192; LDLa; 1.
 DR SMART: SM00202; Tryp_Spc; 1.
 DR PROSITE: PS01209; LDLRA_1; 1.
 DR PROSITE: PS50068; LDLRA_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase; Membrane; Protease; Serine protease; Transmembrane.
 KM HydroLase; Membrane; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 486 AA; 53213 MW; 8618788C5F5A13C CRC64;
 Query Match 25.3%; Score 759.5; DB 2; Length 486;
 Best Local Similarity 34.1%; Pred. No. 3e-34;
 Matches 190; Conservative 77; Mismatches 177; Indels 113; Gaps 19;

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:13:35 ; Search time 50 Seconds

(without alignments)
983.844 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999

Sequence: 1 MERSHGNASPARPPSAGAS.....TEVLPMYKMSHVRFRKS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /EMC_Celerra_SIDS3/ptodata/2/1aa/5 COMB.pap: *
2: /EMC_Celerra_SIDS3/ptodata/2/1aa/6 COMB.pap: *
3: /EMC_Celerra_SIDS3/ptodata/2/1aa/7 COMB.pap: *
4: /EMC_Celerra_SIDS3/ptodata/2/1aa/H COMB.pap: *
5: /EMC_Celerra_SIDS3/ptodata/2/1aa/PCUS COMB.pap: *
6: /EMC_Celerra_SIDS3/ptodata/2/1aa/RE COMB.pap: *
7: /EMC_Celerra_SIDS3/ptodata/2/1aa/backfile1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------------------------|
| 1 | 2999 | 100.0 | 562 | 2 | US-09-879-792-12 Sequence 12, Appl |
| 2 | 2291.5 | 76.4 | 477 | 2 | US-10-177-661-2 Sequence 2, Appl |
| 3 | 2108 | 70.3 | 446 | 2 | US-10-177-661-4 Sequence 4, Appl |
| 4 | 969.5 | 32.3 | 418 | 2 | US-10-177-661-6 Sequence 6, Appl |
| 5 | 741.5 | 24.7 | 492 | 2 | US-09-759-143-932 Sequence 932, App |
| 6 | 741.5 | 24.7 | 492 | 2 | US-10-012-896-932 Sequence 932, App |
| 7 | 741.5 | 24.7 | 492 | 3 | US-10-144-678A-932 Sequence 932, App |
| 8 | 740.5 | 24.7 | 492 | 2 | US-09-343-749-2 Sequence 2, Appl |
| 9 | 740.5 | 24.7 | 492 | 2 | US-09-691-840-2 Sequence 2, Appl |
| 10 | 740.5 | 24.7 | 510 | 2 | US-09-949-016-11074 Sequence 11074, A |
| 11 | 736.5 | 24.6 | 492 | 2 | US-09-685-166A-895 Sequence 895, App |
| 12 | 736.5 | 24.6 | 492 | 2 | US-09-879-792-14 Sequence 14, Appl |
| 13 | 736.5 | 24.6 | 492 | 2 | US-09-679-426-895 Sequence 895, App |
| 14 | 736.5 | 24.6 | 492 | 2 | US-09-759-143-895 Sequence 895, App |
| 15 | 736.5 | 24.6 | 492 | 2 | US-10-012-896-895 Sequence 895, App |
| 16 | 736.5 | 24.6 | 492 | 3 | US-10-144-678A-895 Sequence 895, App |
| 17 | 736.5 | 24.6 | 521 | 2 | US-09-949-016-11081 Sequence 11081, A |
| 18 | 736.5 | 24.6 | 521 | 2 | US-09-949-016-11082 Sequence 11082, A |
| 19 | 736.5 | 24.6 | 521 | 2 | US-09-949-016-11083 Sequence 11083, A |
| 20 | 729.5 | 24.3 | 453 | 2 | US-09-999-833A-69 Sequence 69, Appl |
| 21 | 729.5 | 24.3 | 453 | 2 | US-10-020-445A-69 Sequence 69, Appl |
| 22 | 729.5 | 24.3 | 453 | 2 | US-09-978-188-69 Sequence 69, Appl |
| 23 | 729.5 | 24.3 | 453 | 2 | US-10-017-085A-69 Sequence 69, Appl |
| 24 | 729.5 | 24.3 | 453 | 2 | US-10-145-129A-69 Sequence 69, Appl |
| 25 | 729.5 | 24.3 | 453 | 3 | US-10-013-929A-69 Sequence 69, Appl |
| 26 | 729.5 | 24.3 | 453 | 3 | US-10-013-917A-69 Sequence 69, Appl |

| | | | | | |
|----|-------|------|-----|---|--------------------------------------|
| 27 | 727.5 | 24.3 | 393 | 2 | US-09-759-143-934 Sequence 934, App |
| 28 | 727.5 | 24.3 | 393 | 3 | US-10-012-896-934 Sequence 934, App |
| 29 | 727.5 | 24.3 | 393 | 2 | US-10-144-678A-934 Sequence 934, App |
| 30 | 726 | 24.2 | 454 | 2 | US-09-518-046-2 Sequence 2, Appl |
| 31 | 726 | 24.2 | 454 | 2 | US-09-518-046-2 Sequence 2, Appl |
| 32 | 704 | 23.5 | 423 | 2 | US-09-650-371-2 Sequence 2, Appl |
| 33 | 697.5 | 23.3 | 435 | 2 | US-09-656-002-2 Sequence 2, Appl |
| 34 | 697.5 | 23.3 | 435 | 3 | US-09-607-745-2 Sequence 2, Appl |
| 35 | 697.5 | 23.3 | 437 | 2 | US-10-030-688-2 Sequence 8, Appl |
| 36 | 696.5 | 23.2 | 435 | 2 | US-09-851-588-8 Sequence 6, Appl |
| 37 | 696.5 | 23.2 | 435 | 2 | US-09-008-271A-6 Sequence 6, Appl |
| 38 | 695.5 | 23.2 | 406 | 2 | US-09-968-415-6 Sequence 6, Appl |
| 39 | 694 | 23.1 | 432 | 2 | US-09-851-588-6 Sequence 275, App |
| 40 | 694 | 23.1 | 432 | 2 | US-10-012-231A-275 Sequence 275, App |
| 41 | 694 | 23.1 | 432 | 2 | US-10-015-389A-275 Sequence 275, App |
| 42 | 694 | 23.1 | 432 | 2 | US-10-006-768A-275 Sequence 275, App |
| 43 | 694 | 23.1 | 432 | 2 | US-10-015-671A-275 Sequence 275, App |
| 44 | 694 | 23.1 | 432 | 2 | US-10-015-393A-275 Sequence 275, App |
| 45 | 694 | 23.1 | 432 | 2 | US-10-011-833A-275 Sequence 275, App |
| | | | | | US-10-006-041A-275 Sequence 275, App |

ALIGNMENTS

```
RESULT 1
US-09-879-792-12
; Sequence 12, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-879-792-12
Query Match 100.0%; Score 2999; DB 2; Length 562;
Best Local Similarity 100.0%; Pred. No. 1.9e-204;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MERSHGNASPARPPSAGASPAQSPACTPPGRASPAQSPACTPPGRASPAQAS 60
1 MERSHGNASPARPPSAGASPAQSPACTPPGRASPAQSPACTPPGRASPAQAS 60
61 PAGTPPGASGRASPAQSPARASPALASLSRSSSGSSARASVTTSTPRVLYVAT 120
61 PAGTPPGASGRASPAQSPARASPALASLSRSSSGSSARASVTTSTPRVLYVAT 120
121 EVGAVPIRSSPARAPATRAATRESPTSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
121 EVGAVPIRSSPARAPATRAATRESPTSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
121 EVGAVPIRSSPARAPATRAATRESPTSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
181 FQFQGHGHTGTYKQKRESCPGAVRCGVNDCKLKSDELGVCFPMWDSLTKIYSSSHO 240
181 FQFQGHGHTGTYKQKRESCPGAVRCGVNDCKLKSDELGVCFPMWDSLTKIYSSSHO 240
241 WLPICSSNMNDSYEKTCQQLGFESAHRTTEVAHRDFANFSILRYNSTIOESLHRSCEP 300
```

| | | | |
|----|-----|---|-----|
| Db | 241 | WLPICSSMWNDSYSEKTCQQLQJFESHRKTTYAHNDFANSFSLKRNSTIIQBSLHSECP | 300 |
| Qy | 301 | SQRYSLQCSHCGLPAMTGRIVGALASDSKMPWVSLHFGTHICGTLIDAQWLTAA | 360 |
| Db | 301 | SQRYSLQCSHCGLPAMTGRIVGALASDSKMPWVSLHFGTHICGTLIDAQWLTAA | 360 |
| Qy | 361 | HCFPTREKVLGKMWYVAGTSMULHOLPEAASTAEIININNYDEEDDYDIALMRLSKPLT | 420 |
| Db | 361 | HCFPTREKVLGKMWYVAGTSMULHOLPEAASTAEIININNYDEEDDYDIALMRLSKPLT | 420 |
| Qy | 421 | LSAHIHPACLPMHGQTFSLNFTCMWITGFKETRETDKTSPLREVOVNLIDFKKCDYLV | 480 |
| Db | 421 | LSAHIHPACLPMHGQTFSLNFTCMWITGFKETRETDKTSPLREVOVNLIDFKKCDYLV | 480 |
| Qy | 481 | YDSYITPRMCKAGDIRGGRDSCQGSQSGPLVCBQNNRWYLAGYTSWGTGCGQANRKGVT | 540 |
| Db | 481 | YDSYITPRMCKAGDIRGGRDSCQGSQSGPLVCBQNNRWYLAGYTSWGTGCGQANRKGVT | 540 |
| Qy | 541 | KYTEVLPMIYSQMESEVFRKSS | 562 |
| Db | 541 | KYTEVLPMIYSQMESEVFRKSS | 562 |

```

RESULT 2
US-10-177-661-2
: Sequence 2, Application US/10177661
: Patent No. 6794173
: GENERAL INFORMATION:
: APPLICANT: Anderson, Dirk M.
: APPLICANT: Vitrca, G. Duke
: TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
: FILE REFERENCE: 3256-A
: CURRENT APPLICATION NUMBER: US/10/177,661
: CURRENT FILING DATE: 2002-06-20
: PRIOR APPLICATION NUMBER: US 60/299,606
: PRIOR FILING DATE: 2001-06-20
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO. 2
: LENGTH: 477
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-177-661-2

```

| Query Match | 76.4% | Score 2291.5 | DB 2 | Length 477 |
|-----------------------|--------------|---|---------------|------------|
| Best Local Similarity | 95.7% | Pred. No. 2e-154 | | |
| Matches 424 | Conservative | 4 | Mismatches 12 | Indels 3 |
| | | | Gaps | 1 |
| QY | 120 | TPVGAVPIRSEPARAPATRAATRBSPGSLPKFTWRGQKQPLIGCVLLIALVVSLLI | 179 | |
| | | ::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: | | |
| Db | 38 | SPGISISDSSDPASVSPAAALL--GTSLPKFTWRGQKQPLIGCVLLIALVVSLLI | 94 | |
| QY | 180 | LFOFMQGHITGIRYKQORBSCPHNAVRCQDVNDCKLSDELGCYRPFWDYSLKITYGSSH | 239 | |
| Db | 95 | LFOFMQGHITGIRYKQORBSCPHNAVRCQDVNDCKLSDELGCYRPFWDYSLKITYGSSH | 154 | |
| QY | 240 | QMLPICSSNMWDSYSEKTCQOGLGFSNAHTTEVAHDFPNSFSLILNYSNTIOESLHRSEC | 299 | |
| Db | 155 | QMLPICSSNMWDSYSEKTCQOGLGFSNAHTTEVAHDFPNSFSLILNYSNTIOESLHRSEC | 214 | |
| QY | 300 | PSQRYISLQCSHCGLRAMTGRIVGALASDSKPMQVSLHFGTTHICGGTLIDAOVWLTA | 359 | |
| Db | 215 | PSQRYISLQCSHCGLRAMTGRIVGALASDSKPMQVSLHFGTTHICGGTLIDAOVWLTA | 274 | |
| QY | 360 | AHCFVYTBKULBGMKVYAGTSLNHLPPAAISAEIINSNTYDEEDDYDIALMRSLKPL | 419 | |
| Db | 275 | AHCFVYTBKULBGMKVYAGTSLNHLPPAAISAEIINSNTYDEEDDYDIALMRSLKPL | 334 | |
| QY | 420 | TLSAIIHPACLPMHQGTSLNFCQWLTGEGKTRBTDKTSPLRLRYOVLIDPKKCNDYI | 479 | |
| Db | 335 | TLSAIIHPACLPMHQGTSLNFCQWLTGEGKTRBTDKTSPLRLRYOVLIDPKKCNDYI | 394 | |

| QY | 480 | QY | 480 |
|---|-----|---|-----|
| VDVSYLLTPRMKCGDLRGGRDSCQDSDGPLYCEONRMTYAGYTSMTGCGQGRNKGVY | 539 | VDVSYLLTPRMKCGDLRGGRDSCQDSDGPLYCEONRMTYAGYTSMTGCGQGRNKGVY | 539 |
| VDVSYLLTPRMKCGDLRGGRDSCQDSDGPLYCEONRMTYAGYTSMTGCGQGRNKGVY | 454 | VDVSYLLTPRMKCGDLRGGRDSCQDSDGPLYCEONRMTYAGYTSMTGCGQGRNKGVY | 454 |
| QY | 540 | QY | 540 |
| TKTVEVLPMTYMSGESEVAFRKS | 562 | TKTVEVLPMTYMSGESEVAFRKS | 562 |
| TKTVEVLPMTYMSGESEVAFRKS | 477 | TKTVEVLPMTYMSGESEVAFRKS | 477 |

```

RESULT 3
US-10-177-661-4
: Sequence 4, Application US/10177661
: Patent No. 6794173
: GENERAL INFORMATION:
: APPLICANT: Anderson, Dirk M.
: TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
: FILE REFERENCE: 3256-A
: CURRENT APPLICATION NUMBER: US/10/177,661
: CURRENT FILING DATE: 2002-06-20
: PRIOR APPLICATION NUMBER: US 60/299,606
: PRIOR FILING DATE: 2001-06-20
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 446
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-177-661-4

```

| Query Match | 70.3% | Score 2108 | DB 2108 | Length 446 |
|-----------------------|----------------|---|-----------|------------|
| Best Local Similarity | 88.7% | Pred. No. 1.9e-141 | | |
| Matches 393 | Conservative 4 | Mismatches 12 | Indels 34 | Gaps 2 |
| QY | 120 | TPYGAVPINSSPPARSAAPATRAATRSBPSTSLPFTWRBSQOKPLIGCYLLIALIVSLIT | 179 | |
| DB | 38 | SPLTSSISSSPASPYSPAAALUL---GTSLPKFTWRBQOKPLIGCYLLIALIVSLIT | 94 | |
| QY | 180 | LTPQMOGHGIRIKREORRESCPGHAYRCQGVYDCKLKSDELCAVFPMDKSLTKIYSSSH | 239 | |
| DB | 95 | LTPQMOGHGIRIKREORRESCPGHAYRCQGVYDCKLKSDELCAVFPMDKSLTKIYSSSH | 154 | |
| QY | 240 | QWLPICSSMWNDSYSSEKTCQQLGFESAHRTTEVAHRDFANSFSLRYNSTIIOBSLHRSEC | 299 | |
| DB | 155 | QWLPICSSMWNDSYSSEKTCQQLGFES-----RSEC | 183 | |
| QY | 300 | PSORYSLIOCSHCGRAMTGRIVGALASDSKMPQVSLHRTTHICGGTLIDAWULTA | 359 | |
| DB | 184 | PSORYSLIOCSHCGRAMTGRIVGALASDSKMPQVSLHRTTHICGGTLIDAWULTA | 243 | |
| QY | 360 | AHCFVYTRKVLKGMKRVAGTSLNLQLPKASIAEIIINSYTDDEDDYDIALMRSLKPL | 419 | |
| DB | 244 | AHCFVYTRKVLKGMKRVAGTSLNLQLPKASIAEIIINSYTDDEDDYDIALMRSLKPL | 303 | |
| QY | 420 | TLSAHHHPACLPMHQGTFSLNTECMITPGKTRTDDTSPFLREVOYNLIDPKKNDYL | 479 | |
| DB | 304 | TLSAHHHPACLPMHQGTFSLNTECMITPGKTRTDDTSPFLREVOYNLIDPKKNDYL | 363 | |
| QY | 480 | VYDSYLTTPRMKCGADLRGRDSCQGDSDGAPLVCBONNRWYLAGVTSMTGCGGGRNKPGVY | 539 | |
| DB | 364 | VYDSYLTTPRMKCGADLRGRDSCQGDSDGAPLVCBONNRWYLAGVTSMTGCGGGRNKPGVY | 423 | |
| QY | 540 | TKYTEVLPMTYSKMESEVAFKRS | 562 | |
| DB | 424 | TKYTEVLPMTYSKMESEVAFKRS | 446 | |

```

RESULT 4
US-10-177-661-6
; Sequence 6, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.

```

APPLICANT: Vitca, G. Duke
TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
FILE REFERENCE: 3256-A
CURRENT APPLICATION NUMBER: US/10/177,661
CURRENT FILING DATE: 2002-06-20
PRIORITY APPLICATION NUMBER: US 60/299,606
PRIORITY FILING DATE: 2001-06-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 418
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Consensus sequence
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)-(418)
OTHER INFORMATION: Xaa = unknown
US-10-177-661-6

Query Match 32.3%; Score 969.5; DB 2; Length 418;
Best Local Similarity 50.2%; Pred. No. 8.2e-61;
Matches 208; Conservative 27; Mismatches 130; Indels 49; Gaps 12;

OY 198 SCFPAHRCGVDVCKLSDCLG-----VRPMDKSLIKYSSGSHQMLPTCS 246
DB 1 SCIFPSKXKCDVXDCPEXGDEXKCVXXXXXGPPXXVRLYGDXXLLQVYSSXXXXXKVPVCS 60
OY 247 SNMDSYSEKTCQOLGFESAHRTTE--VAHRDFANSFSLR-----YNSTIOES 293
DB 61 DNMESEYXXACXQMGXXSAYYSEKXXKXGANSFPAKLVSPXNLLXHXYSIXXX 120
OY 294 LHRSE--CPGQRYISLQCS--HCGLR---AMTRIVGALASDKPMQVSLHFGT-T 343
DB 121 LKRSXXKCPGXXVYSLQCSXQDCGRLNAXXMTSRIVGXXASXGKMPQVSLQXXGV 180
OY 344 HICGGLIDAQWVLTAAHCFVTRKVLKGVKAVAG--TSNLIHQPEAA--SIAEIINS 399
DB 181 HLCGSLIXPFWVLTAAHCVGRXXKPLKGMXVFAGILTXSLHXXPKAXXXVKKIHP 240
OY 400 NY-----TDEEDDDIALMRSLKPLTSAHHPACLPFHGOTFSLNCTWITGFKTRE 453
DB 241 NYXXXXXXKXKXNDIAMLKSLKPLTXDYOPLVCLNPGOXLXPGTTCIXMGATXE 300
OY 454 TDKTSPFLREVOVNLIDFKCNDYLVYDSYLTFRMCAQDLRGGRDSCQDGGPLVCE 513
DB 301 XXGTSPTVLEAXVPLIDNKKXNSYXVYDNXITPRMTCAGVLEGGVDSQDGGPLVCE 360
OY 514 ---QNNRWYLAGVTSWGTGC-GQNRKGVYTKTEVLPWYISKMESEVRRKS 562
DB 361 XXXXQNNRWMLXGXTSMGXGCAKAKPKGVYTXVYFLXWYSQXXEXRRFKS 414

RESULT 5
US-09-759-143-932
Sequence 932, Application US/09759143
Patent No. 6800746
GENERAL INFORMATION:

APPLICANT: Xu, Jiaoguan
APPLICANT: Dillon, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqiu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hegler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-932

Query Match 24.7%; Score 741.5; DB 2; Length 492;
Best Local Similarity 32.3%; Pred. No. 1.4e-44;
Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

OY 62 AGTPRGRA-----SPGRASPAQAS-----PARASPALASLRSSGSSRSRS 104
DB 5 SGSPPAIGPYENHGYOPENPYPAQPTVTVVEHPAQYYP----- 46
OY 105 ASVTPSPRYLVATPVGAVPPIRSSPARSAPATRAITRESPTSLPKFTWREGOKLP- 163
DB 47 -----SPVQYAPRVLTQASNPVCTQPKSPSGTVCSTK-----KCALCIT 88
OY 164 --ICGVLLIALLVSLILFQWQ-----GHTGIRKEQRESCPKAIVRCGVVDCXK 215
DB 89 LTLGTFVGAALAGL-----WRFMGSKCSNSGIEC-DSSGTCINPSMCDGVSHCPG 142
OY 216 SDELGCYRFMDKSLIKYSSGSHQMLPTCSNNNDYSEKTCQOLGFESAHRTTEVAHR 275
DB 143 EDENRCVRLYGSNFTLOYSSQKRSKMHVCCDDNENYGRACDMDGKKNFYSSQ-GIV 201
OY 276 DFANSFSLRNTST-----IOESLHRSE-CPGQRYISLQCSHGLR---AMTRIVGAL 326
DB 202 DSGSTSPKMLNTSAGVNDIYKLYHSDACSKRAVLSRLACGVNLSNRSQSRIVGES 261
OY 327 ASDSKPMQVSLHFGTTHICGGLIDAQWVLTAAHCFVTRKVLKGVKAVAG--TSNLI 383
DB 262 ALPAMPQVSLHQNVAHVCSSITPEWITVAHCV-----EKPLNPMWHTAFAGILRQ 317
OY 384 HOLPEAA--SIAEIINSNTYDEEDDDIALMRSLKPLTSAHHPACLPFHGOTFSLNE 441
DB 318 SFMFYAGYQYKESISHENYDSKTKNDIALMKLOKPLTFNDLVKPVCLPMPGMLOPEQ 377
OY 442 TCWITGFKTRETDKTSPLREVOVNLIDFKCNDYLVYDSYLTFRMCAQDLRGGRDSC 501
DB 378 LCMISGMKATIE-KGTSEVINAAKVLLIETORCNSRYVDNLITPAMICAGFLOGNVD 436
OY 502 CQDGGGPLVCEQNNRWYLAGVTSWGTGCGQNRKGVYTKTEVLPWYISKMESE 556
DB 437 CQDGGGPLVTSKXNIMWLIQDTSMGSCAKAYRPGVYGNVFTDVIYRMRAD 491

RESULT 6
US-10-012-896-932
Sequence 932, Application US/10012896
Patent No. 6943236
GENERAL INFORMATION:

APPLICANT: Xu, Jiaoguan
APPLICANT: Dillon, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqiu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick

APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Mantanabe, Yoshinhiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-932

Query Match 24.7%; Score 741.5; DB 2; Length 492;
Best Local Similarity 32.3%; Pred. No. 1.4e-44;
Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

62 AGTPPGRA-----SPGRASPAQAS-----PARASPALASLSRSSSGSSARS 104
5 SGSPPAIPYENHGYOPENPYPAQPTVPTVEVHPAQYPT-----KALCIT 88
105 ASVTSPTRVLYVATPVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGOKOLPL- 163
47 -----SPVQYAPRVLTOASNPVCTQPKSPSGVCTSKT-----KALCIT 88
164 --ICVLLIALLVSLIILFOFWQ-----GHTGIRYKQRESQCPKAVRCQGVVDCIK 215
89 LTLGTFVLGALAAAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPSNMCDSVSHCPGQ 142
216 SDELCGVRPMDKSLIKTYSGSSHOWLPIGSSNNDSVSEKTCQQLGFGESAHRTTEVAHR 275
143 EDENRCVRLYGSNFILOYSSQKSMHPCVODDNNENNGRAACRDMGKKNFYSSQ-GIV 201
276 DFANSFSLIRYNT-----IQESLHRSE-CPQRYISIQCSHCGLR---AMTGRIVGAL 326
202 DDSGSTPMKNTSAGVNDIYKQLYHSDACSSKAVVSLRCLACGVNLNSSQSRIVGES 261
327 ASDSKPMQVSLHFGTTHICGTLIDAQWVLTAAHCFVTRKYLEG--WKVYAGTSL 383
262 ALPGAMPQVSLHGVNVAVCGSIIITPEMITAAHCV---EKPLNPMHTAFAGILRQ 317
384 HOLPEAA--STAETIINSNTDEEDVDIALMRLSKPLTSAHHPACLPNHGOTFSINE 441
318 SFMFYGAQYQVEKISHPNYDSKTKNNDIALMKLOKPLTFNDLVKPCVLPMPGMLQPEQ 377
442 TCWITGFKTRETDDKTSPLREVQVNLIDFKKCNLDVLYVDSYLTTPMMACAGDLRGDS 501
378 LCMWISGMGATEE-KGKTSBVANAQVLLIETQRCNSRYVDNLTTPMICAQGFQGNVDS 436
502 CGDSGGPLVCEONNRWYLAGVTSWGTGCGGRNKRPGVYTKTEVLPWYISKMESE 556
437 CGDSGGPLVSKNNIMWLIGDTSWGSQCAKAYRGVYGNVTFDWTYRQMRAD 491

RESULT 7
US-10-144-678A-932
Sequence 932, Application US/10144678A
Patent No. 7033827
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Bassols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Mantanabe, Yoshinhiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144,678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-678A-932

Query Match 24.7%; Score 741.5; DB 3; Length 492;
Best Local Similarity 32.3%; Pred. No. 1.4e-44;
Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

62 AGTPPGRA-----SPGRASPAQAS-----PARASPALASLSRSSSGSSARS 104
5 SGSPPAIPYENHGYOPENPYPAQPTVPTVEVHPAQYPT-----KALCIT 88
105 ASVTSPTRVLYVATPVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGOKOLPL- 163
47 -----SPVQYAPRVLTOASNPVCTQPKSPSGVCTSKT-----KALCIT 88
164 --ICVLLIALLVSLIILFOFWQ-----GHTGIRYKQRESQCPKAVRCQGVVDCIK 215
89 LTLGTFVLGALAAAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPSNMCDSVSHCPGQ 142
216 SDELCGVRPMDKSLIKTYSGSSHOWLPIGSSNNDSVSEKTCQQLGFGESAHRTTEVAHR 275
143 EDENRCVRLYGSNFILOYSSQKSMHPCVODDNNENNGRAACRDMGKKNFYSSQ-GIV 201
276 DFANSFSLIRYNT-----IQESLHRSE-CPQRYISIQCSHCGLR---AMTGRIVGAL 326
202 DDSGSTPMKNTSAGVNDIYKQLYHSDACSSKAVVSLRCLACGVNLNSSQSRIVGES 261
327 ASDSKPMQVSLHFGTTHICGTLIDAQWVLTAAHCFVTRKYLEG--WKVYAGTSL 383
262 ALPGAMPQVSLHGVNVAVCGSIIITPEMITAAHCV---EKPLNPMHTAFAGILRQ 317
384 HOLPEAA--STAETIINSNTDEEDVDIALMRLSKPLTSAHHPACLPNHGOTFSINE 441
318 SFMFYGAQYQVEKISHPNYDSKTKNNDIALMKLOKPLTFNDLVKPCVLPMPGMLQPEQ 377
442 TCWITGFKTRETDDKTSPLREVQVNLIDFKKCNLDVLYVDSYLTTPMMACAGDLRGDS 501
378 LCMWISGMGATEE-KGKTSBVANAQVLLIETQRCNSRYVDNLTTPMICAQGFQGNVDS 436
502 CGDSGGPLVCEONNRWYLAGVTSWGTGCGGRNKRPGVYTKTEVLPWYISKMESE 556

Db 437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

RESULT 8

US-09-342-749-2

; Sequence 2, Application US/09342749

; Patent No. 6166194

; GENERAL INFORMATION:

; APPLICANT: Wong, Alexander K.C.

; APPLICANT: Tavtigian, Sean V.

; APPLICANT: Teng, David H.-F.

; APPLICANT: Myriad Genetics, Inc.

; TITLE OF INVENTION: TPRS2 is a Tumor Suppressor

; FILE REFERENCE: 2318-202

; CURRENT APPLICATION NUMBER: US/09/342,749

; CURRENT FILING DATE: 1999-06-29

; EARLIER APPLICATION NUMBER: US 60/091,044

; EARLIER FILING DATE: 1998-06-29

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 492

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-342-749-2

Query Match 24.7%; Score 740.5; DB 2; Length 492;

Best Local Similarity 32.3%; Pred. No. 1.7e-44;

Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

62 AGTPPGR-----SPGRSPAAS-----PARASPALASLRSSSGSSSARS 104

5 SGSPPAIGPYENHGYOENPYPAQPTVPTVEVHPAQYR----- 46

105 ASVTTPRYLVATPVGAVPIRSSPARSAPATPATRESPTSLPKFTWEGOKLP- 163

47 -----SPVQYAPRVLTQASNPVCTQPKSPSGTCTKT-----KALCIT 88

164 --ICVLLILALVSLILFQFWQ-----GHTGIRYQESRCPKHAVRCDGVDCIK 215

89 LTLGTFVLGALLAGLL-----WKFMGSKCSNSGIEC--DSGCTCINPSNWCDCVSHCPGG 142

216 SDELCVAFDMDKSLIKYSSGSHOVLPICSSNMNDSYSEKTCOOLGFESAHTREVAHR 275

143 EDENRCVALLYPNFTLQVYSSQKSMHPCVCDNMENNGRAACRMGYKNNFYSSQ-GIV 201

276 DPANFSILRYNST-----IOESLRSE-CPSORYISLCQSHGRL--AMTGRIVGGAL 326

202 DSGSTSMKLNKNTSAGNVDIYKLYHSDACSKAVSLRCLACGVNLNLSRQSRIVGGS 261

327 ASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLG--WKVYAGTNTL 383

262 ALPGAMPQVSLHYQNVHVCSSITLPEMIVTAHCV-----EKPLNPMHMTAFAGILRQ 317

384 HOLPEAA--STAEIINSNTYDEEDDYDIALMRSLKPLTSLAHIPACLPNMGOTFSINE 441

318 SFMFYAGYQVEKVIHSPNYDSKTKNDIALMKLOKPLTFNDLVKVCPLPNEGMLQEQ 377

442 TCWITGFKTRTDDKTSPLREVOVNLIDFKKCNLYLVYDSYLTPRMNCAGDLRGDS 501

378 LQWISGKATEE-KKTSSEVLNAKAVLLIETORCNSRYVYDNLITPAMICAGFLQGNDS 436

502 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

437 CQGSQGGPLVTSKNNIMWLIDTWSGSCAKAYRGGVGNVWFTDWIRQMRAD 491

```

? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 11074
? LENGTH: 510
? TYPE: PR1
? ORGANISM: Human
US-09-949-016-11074

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 24.7%; | Score 740.5; | DB 2; | Length 510; |
| Best Local Similarity | 32.3%; | Pred. No. 1,8e-44; | | |
| Matches 173; | Conservative 85; | Mismatches 189; | Indels 88; | Gaps 16; |

[illegible]

RESULT 11
US-09-685-166A-895
Sequence 895 Application US/09685166A
Patent No.6630305
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlockner, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel

```

1  APPLICANT: Wang, Aifun
2  APPLICANT: Skeiky, Yasir A.W.
3  APPLICANT: Hopley, William
4  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
5  TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
6  FILE REFERENCE: 210121.427C21
7  CURRENT APPLICATION NUMBER: US/09/685,166A
8  CURRENT FILING DATE: 2000-10-10
9  NUMBER OF SEQ. ID NOS: 898
10 SOFTWARE: PastedSeq for Windows Version 3.0
11 SEQ ID NO 895
12 LENGTH: 492
13 TYPE: PRT
14 ORGANISM: Homo sapiens
15 US-09-685-166A-895

```

| | | | | |
|-----------------------|--------------|--------------------|------------|-------------|
| Query Match | 24.6%; | Score 736.5; | DB 2; | Length 492; |
| Best Local Similarity | 32.4%; | Pred. No. 3.2e-44; | | |
| Matches 173; | Conservative | 84; | Mismatches | 189; |
| | | | Indels | 88; |
| | | | Gaps | 16; |

[illegible]

```

RESULT 12
US-09-879-792-14
; Sequence 14, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Proctasease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13

```

; PRIOR APPLICATION NUMBER: US 60/283,648
 ; PRIOR FILING DATE: 2001-04-16 (Docket No. 6734006 LIO-81-WO)
 ; PRIOR APPLICATION NUMBER: PCT
 ; PRIOR FILING DATE: 2001-06-12
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-879-792-14

Query Match 24.6%; Score 736.5; DB 2; Length 492;
 Best Local Similarity 32.4%; Pred. No. 3.2e-44;
 Matches 173; Conservative 84; Mismatches 189; Indels 88; Gaps 16;

QY 62 AGTPPGR-----SPGRASPAQAS-----PARASPALASLSRSSSGRSSARS 104
 Db 5 SGSPPAIGPYENHGYQENPYPAQPTVPTVEVHPAQYV----- 46
 QY 105 ASVTPSPRYLVNATPVGAVPIRSSPARSAPATRATESPOTSLPKFTWREGOKLP- 163
 Db 47 -----SPVQYAPRVLTQASNPVCTQPKPSGVTCTSKT-----KCALCIT 88
 QY 164 --IGCVLLIALVLSLILFQFWQ-----GHTGIRYKQRESCEPKAVRCDCGVDCIK 215
 Db 89 LTLGFLVGLAALAGL-----WKFMGSKCNSGIEC-DSSGTCINPNCDCGVSHCPG 142
 QY 216 SDELGVCFWMDKSLIKYSGSSHOWLPICSSNNNDYSEKTCQOLGFEASHRTTEVAHR 275
 Db 143 EDENRCVLYPNPFLQWYSSQKSMHVCQDDMNENYGRAACRMGKKNFYSSQ-CIV 201
 QY 276 DFANSFSLRYNST-----IOESLHRS-CPSORYISIQCSHGRL--AMTGRIVGAL 326
 Db 202 DDSGSTFMKNTSAGNDYIKKLYHSDACSSKAVVSLRCLACGVNLSSRQSRIVGES 261
 QY 327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTEKVEG--WKVYAGTSL 383
 Db 262 ALPGAMPQVSLHVNVAHVCSSITTPEMIVTAHCV---EKPLNNPMHTAFAGILRQ 317
 QY 384 HOLPEAA--STAEIINSNYDEEDDYIALMRSLKPTLSAHIHAPCLPMHGQFSLNE 441
 Db 318 SFMFYAGYQVQKISHPNYDSKTKNDIALMKLOKPLTFNDLVKPVCLPMPGMLOPEQ 377
 QY 442 TCWITGFKETRETDDKTPFLREVQVNLIDPKKNDYLVYDSYLPFPMACADLRGDS 501
 Db 378 LCMISGMWATEE-KGKTSEVLAARVLLIETORCNSRYVDNLITPAMICAGFLOGNDS 436
 QY 502 CQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRNKPQVYTKTEVLPWITYSKMES 555
 Db 437 CQDSSGGLVTSNNIMWLIGDTSWGSACAKAYRPGVYGVMVFTDWIYRQKA 490

RESULT 13
 US-09-679-426-895
 ; Sequence 895, Application US/09679426
 ; Patent No. 6759515
 ; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Devin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqun
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stoik, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.
 ; APPLICANT: Carter, Derrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yaser A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C20
 ; CURRENT FILING DATE: US/09/679,426
 ; NUMBER OF SEQ ID NOS: 2000-10-02
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 895
 ; LENGTH: 492
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-679-426-895

Query Match 24.6%; Score 736.5; DB 2; Length 492;
 Best Local Similarity 32.4%; Pred. No. 3.2e-44;
 Matches 173; Conservative 84; Mismatches 189; Indels 88; Gaps 16;

QY 62 AGTPPGR-----SPGRASPAQAS-----PARASPALASLSRSSSGRSSARS 104
 Db 5 SGSPPAIGPYENHGYQENPYPAQPTVPTVEVHPAQYV----- 46
 QY 105 ASVTPSPRYLVNATPVGAVPIRSSPARSAPATRATESPOTSLPKFTWREGOKLP- 163
 Db 47 -----SPVQYAPRVLTQASNPVCTQPKPSGVTCTSKT-----KCALCIT 88
 QY 164 --IGCVLLIALVLSLILFQFWQ-----GHTGIRYKQRESCEPKAVRCDCGVDCIK 215
 Db 89 LTLGFLVGLAALAGL-----WKFMGSKCNSGIEC-DSSGTCINPNCDCGVSHCPG 142
 QY 216 SDELGVCFWMDKSLIKYSGSSHOWLPICSSNNNDYSEKTCQOLGFEASHRTTEVAHR 275
 Db 143 EDENRCVLYPNPFLQWYSSQKSMHVCQDDMNENYGRAACRMGKKNFYSSQ-CIV 201
 QY 276 DFANSFSLRYNST-----IOESLHRS-CPSORYISIQCSHGRL--AMTGRIVGAL 326
 Db 202 DDSGSTFMKNTSAGNDYIKKLYHSDACSSKAVVSLRCLACGVNLSSRQSRIVGES 261
 QY 327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTEKVEG--WKVYAGTSL 383
 Db 262 ALPGAMPQVSLHVNVAHVCSSITTPEMIVTAHCV---EKPLNNPMHTAFAGILRQ 317
 QY 384 HOLPEAA--STAEIINSNYDEEDDYIALMRSLKPTLSAHIHAPCLPMHGQFSLNE 441
 Db 318 SFMFYAGYQVQKISHPNYDSKTKNDIALMKLOKPLTFNDLVKPVCLPMPGMLOPEQ 377
 QY 442 TCWITGFKETRETDDKTPFLREVQVNLIDPKKNDYLVYDSYLPFPMACADLRGDS 501
 Db 378 LCMISGMWATEE-KGKTSEVLAARVLLIETORCNSRYVDNLITPAMICAGFLOGNDS 436
 QY 502 CQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRNKPQVYTKTEVLPWITYSKMES 555
 Db 437 CQDSSGGLVTSNNIMWLIGDTSWGSACAKAYRPGVYGVMVFTDWIYRQKA 490

RESULT 14
 US-09-759-143-895
 ; Sequence 895, Application US/09759143
 ; Patent No. 6800746
 ; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Devin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqun
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stoik, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.

```

? APPLICANT: Carter, Darrick
? APPLICANT: Li, Samuel
? APPLICANT: Wang, Ajun
? APPLICANT: Skeiky, Yaser A.W.
? APPLICANT: Hepler, William
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
? TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
? FILE REFERENCE: 210121.427623
? CURRENT APPLICATION NUMBER: US/09/759,143
? CURRENT FILING DATE: 2001-01-12
? NUMBER OF SEQ ID NOS: 934
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 895
? LENGTH: 492
? TYPE: PR1
? ORGANISM: Homo sapiens
US-09-759-143-895

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 24.6%; | Score 736.5; | DB 2; | Length 492; |
| Best Local Similarity | 32.4%; | Pred. No. 3.2e-44; | | |
| Matches 173; | Conservative 84; | Mismatches 189; | Indels 88; | Gaps 16 |

| | | | |
|----|-----|---|-----|
| QY | 62 | AGTGPGR-----SGRASPAAS-----PARASALASISSSGRSSARS | 104 |
| Db | 5 | SGSPPAIGPYENHCYOPENPRPADPTVPTVYVEHQAQYV----- | 46 |
| QY | 105 | ASVTTSPRVYLVLRATPVGAVPIRSSPARASAPATATBSPOTSILPKFTWREGQOLPL- | 163 |
| Db | 47 | -----SPPOVAPRVLTOASNPVYCTOPKSPSGVCTST-----KRALCT | 88 |
| QY | 164 | --IGCVLLILALVSVLILFQFWQ-----GHTGYRKEQRESCPEKHAVCDDVDCXKL | 215 |
| Db | 89 | LTLGFLVGAALAAQL-----WKRMGSKCSNSGLEC-DSSGTCINPSWCDGVCHPCPG | 142 |
| QY | 216 | SDELCAVRFMDKSLKTIYSGSSHOWLPICSSNNMDSYSEKTCQOLGFEASARHTEVAHR | 275 |
| Db | 143 | EDENRCVLYGPNFLQWYSSQRKSMHPYCCODMVENYGRACRDMYKNNFYSSQ-GIV | 201 |
| QY | 276 | DEANSFSLIRYNST-----IOESLRSE-CPSORSTISQCSHGRL--AMGRIYGAAL | 326 |
| Db | 202 | DDSGSTSMKANTSAANDIYKLYHSDACSAVSLRCLACGVNLNSRBSRIYGGSS | 261 |
| QY | 327 | ASDSKMPQVSLHFGTTHICGGTLIDAOVLTAAHCFVTRKVLBG--WKVYAGTSL | 383 |
| Db | 262 | ALPGAMPQVSLHVONVHVCSSGSLITPEWITYAACV-----EKLNNPWHMTAFAGILRO | 317 |
| QY | 384 | HQLPEAA--STAIIINSNYTDEBDYDIALMRLEKPLTSLASHIPACLPYMGQGFSLNE | 441 |
| Db | 318 | SFMRYGAGYQVOKVLSHPNYDSKTDNDIALMKLOKPLTENDLVFVCLIPNFGMLQEPQ | 377 |
| QY | 442 | TCWITGFGKTRTDDKTSPLREVOVNLIDFKCNDYLVYDSYLTTPRMWCAGDLRGARDS | 501 |
| Db | 378 | LCWISGMATEE--KCKTSEVLNAAKVLILETORCNSRYVDNLTIPAMICAGFLQGNVDS | 436 |
| QY | 502 | COGSGGGLVCEONNRWYLAGVTSKGTGQGNKRGVYTKYTEVLPMYISKES | 555 |
| Db | 437 | COGSGGGLVYNNNIWMLIGTSSKSGGAKARYRGVYGNVWVFTDMITRQKA | 490 |

RESULT 15
US-10-012-696-695
Sequence 895 Application US/100128996
Patent No 6643236
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Devin C.
APPLICANT: Harlocker, Jennifer L.
APPLICANT: Hitchcock, Susan L.
APPLICANT: Jiang, Yudu
APPLICANT: Katos, Michael D.
APPLICANT: Reiter, Marc W.
APPLICANT: Stok, John A.
APPLICANT: Day, Craig H.

```

APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Wantanabe, Yoshitiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 895
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-895

```

| | | | | |
|-----------------------|-------|-------------------|------|----------------|
| Query Match | 24.6% | Score 736.5 | DB 2 | Length 492 |
| Best Local Similarity | 32.4% | Pred. No. 3 2e-44 | | |
| Best Local | 173 | Conservative | 84 | Mismatches 189 |
| | | | | Indels 88 |
| | | | | Gaps 16 |

```

QY 62 AGTPRGA-----SPGRSPQAO-----PARAPALASIRSSSGRSSARS 104
Db 5 SGRPALGPYYENHGYORENPYRQPTVVYVEVHRAQYR----- 46
QY 105 ASVTTSPTRYVLYRATRVGAVPRIRSSPARSAPATRAITRESQTSLPKTFREGQKQLE- 163
Db 47 -----SVPQAPRPFVLTOASNPVVCTOPKPSQGVCTSTK-----KVALCIT 88
QY 164 --IGCVLLTALVYSLIILPQFWO-----GHTGRYKQRESCPKHAVRCDGVUDKLY 215
Db 89 LTLGTFLVGAALAGLL-----WFPMSKCSNSGIEC-DSSGTCINPNMCDGVSHCCG 142
QY 216 SDELGCYRFDWDSKLTKIYSSSHQWLPICSSNNDSYSEKTCQOLGFSAHRTTEVAHR 275
Db 143 EDERNCRYLGPNFILQWYSSORKSMHPVCCDDMNENBGRACRDMGXONFYSSQ-GIV 201
QY 276 DFANFSILRRNSF-----IOESLHRSE-CPQORYISLOCSHGDR---AANTGRVGGAL 326
Db 202 DDSGTSFPMKINTAGVNDIYKLIYHSDAGCSKAVSILRCLACGVNLNSSQSRLVGGBS 261
QY 327 ASDSKWPMQVSLHFGTTHICGGLTIIDQAWLTAAHCFVTRKXVLEG--WKVAGTSLN 383
Db 262 ALPAPAMQVSLHAYONVHVCSSIIITPEWITYAAHCV-----EKLNNPMHWTAFAGILRO 317
QY 384 HQLEPA--SLAEIINSNYTDEDDVDVIALMRSKPTLSAHIHAPLCPMHCQFSLNE 441
Db 318 SFMYGYAGYQVQAKISHPNYDSKKNNDIALMKQKPLTFNDLYKVCULPYPGMMLQREBQ 377
QY 442 TCMITGEGKTRTDDTKSFFLREYQVNLIDFKCNDVLYVUDSYLTPRMACAGDLRGADS 501
Db 378 LCMISGMGATBE-KGKTSVLNAKAVLLIETQRCNSRYVDNLITPRACIGAGFIQGVNDS 436
QY 502 CQSGSGRPVLEQNNRMYTAGVTSWNGCGCGGRNRPGVYTKYTEVLPWYISAGS 555
Db 437 CQGSQSGRPVLTSSNNIMWLIGDTWSGSCAAAPRYGVYGNMVFETDMLYIQOKA 490

```

Search completed: September 16, 2006, 03:15:05
Job time : 52 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:25:36 ; Search time 181 Seconds
(without alignments)
1438.269 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999

Sequence: 1 MERSHGNASPARTPSAGAS.....TEVLPWYKMSREVRPKS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celextra_sids3/ptodata/2/pubppaa/us07_PUBCOMB.pep:*
- 2: /EMC_Celextra_sids3/ptodata/2/pubppaa/us08_PUBCOMB.pep:*
- 3: /EMC_Celextra_sids3/ptodata/2/pubppaa/us09_PUBCOMB.pep:*
- 4: /EMC_Celextra_sids3/ptodata/2/pubppaa/us10_PUBCOMB.pep:*
- 5: /EMC_Celextra_sids3/ptodata/2/pubppaa/us10B_PUBCOMB.pep:*
- 6: /EMC_Celextra_sids3/ptodata/2/pubppaa/us11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------------------------|
| 1 | 2999 | 100.0 | 562 | 3 | US-09-879-792-12 Sequence 12, Appl |
| 2 | 2999 | 100.0 | 562 | 4 | US-10-806-370-12 Sequence 12, Appl |
| 3 | 2999 | 100.0 | 569 | 4 | US-10-428-275-130 Sequence 130, App |
| 4 | 2991 | 99.7 | 562 | 4 | US-10-156-214A-24 Sequence 24, Appl |
| 5 | 2991 | 99.7 | 562 | 5 | US-10-916-758-94 Sequence 94, Appl |
| 6 | 2963 | 98.8 | 688 | 4 | US-10-156-214A-26 Sequence 26, Appl |
| 7 | 2963 | 98.8 | 688 | 5 | US-10-916-758-2 Sequence 2, Appl |
| 8 | 2956 | 98.6 | 581 | 4 | US-10-353-690-100 Sequence 100, App |
| 9 | 2948.5 | 98.3 | 586 | 4 | US-10-428-275-104 Sequence 104, App |
| 10 | 2948.5 | 98.3 | 586 | 4 | US-10-428-275-124 Sequence 124, App |
| 11 | 2948.5 | 98.3 | 593 | 4 | US-10-428-275-120 Sequence 120, App |
| 12 | 2943.5 | 98.1 | 586 | 4 | US-10-428-275-142 Sequence 142, App |
| 13 | 2942.5 | 98.1 | 586 | 4 | US-10-428-275-140 Sequence 140, App |
| 14 | 2775.5 | 92.5 | 537 | 3 | US-09-888-615-104 Sequence 104, App |
| 15 | 2775.5 | 92.5 | 537 | 4 | US-10-428-275-128 Sequence 128, App |
| 16 | 2775.5 | 92.5 | 537 | 6 | US-11-037-243-104 Sequence 104, App |
| 17 | 2686 | 89.6 | 542 | 4 | US-10-428-275-122 Sequence 122, App |
| 18 | 2291.5 | 76.4 | 477 | 4 | US-10-177-661-2 Sequence 2, Appl |
| 19 | 2291.5 | 76.4 | 477 | 5 | US-09-898-837A-15 Sequence 15, Appl |
| 20 | 2287 | 76.3 | 421 | 3 | US-09-898-837A-15 Sequence 15, Appl |
| 21 | 2287 | 76.3 | 421 | 4 | US-10-428-275-132 Sequence 132, App |
| 22 | 2219.5 | 74.0 | 486 | 4 | US-10-274-639-8 Sequence 8, Appl |
| 23 | 2219.5 | 74.0 | 486 | 4 | US-10-333-574-8 Sequence 8, Appl |
| 24 | 2130.5 | 71.0 | 412 | 4 | US-10-428-275-144 Sequence 144, App |
| 25 | 2108 | 70.3 | 446 | 4 | US-10-177-661-4 Sequence 4, Appl |
| 26 | 2108 | 70.3 | 446 | 5 | US-10-910-507-4 Sequence 4, Appl |
| 27 | 2097 | 69.9 | 382 | 4 | US-10-428-275-134 Sequence 134, App |

| | | | | | |
|----|--------|------|-----|---|-------------------------------------|
| 28 | 2059 | 68.7 | 406 | 4 | US-10-428-275-126 Sequence 126, App |
| 29 | 2059 | 68.7 | 406 | 4 | US-10-428-275-138 Sequence 138, App |
| 30 | 2026 | 67.6 | 401 | 4 | US-10-428-275-112 Sequence 112, App |
| 31 | 2016 | 67.2 | 401 | 4 | US-10-428-275-116 Sequence 116, App |
| 32 | 2012 | 67.1 | 401 | 4 | US-10-428-275-118 Sequence 118, App |
| 33 | 1842.5 | 61.4 | 370 | 4 | US-10-428-275-114 Sequence 114, App |
| 34 | 1839.5 | 61.3 | 367 | 4 | US-10-428-275-110 Sequence 110, App |
| 35 | 1311 | 43.7 | 296 | 3 | US-09-804-156-30 Sequence 30, Appl |
| 36 | 1311 | 43.7 | 296 | 4 | US-10-067-761-30 Sequence 30, Appl |
| 37 | 1311 | 43.7 | 296 | 4 | US-10-319-519-30 Sequence 30, Appl |
| 38 | 1311 | 43.7 | 372 | 3 | US-09-804-156-15 Sequence 15, Appl |
| 39 | 1311 | 43.7 | 372 | 3 | US-09-946-633-7 Sequence 7, Appl |
| 40 | 1311 | 43.7 | 372 | 4 | US-10-125-459-7 Sequence 7, Appl |
| 41 | 1311 | 43.7 | 372 | 4 | US-10-067-761-15 Sequence 15, Appl |
| 42 | 1311 | 43.7 | 372 | 4 | US-10-319-519-15 Sequence 15, Appl |
| 43 | 1270 | 42.3 | 239 | 4 | US-10-428-275-106 Sequence 106, App |
| 44 | 1270 | 42.3 | 239 | 4 | US-10-428-275-108 Sequence 108, App |
| 45 | 1265 | 42.2 | 231 | 4 | US-10-428-275-136 Sequence 136, App |

ALIGNMENTS

```
RESULT 1
US-09-879-792-12
; Sequence 12, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:
; APPLICANT: Xieo, Yonghong
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-879-792-12
Query Match 100.0%; Score 2999; DB 3; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MERSHGNASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
1 MERSHGNASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
61 PAGTPGRASGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
61 PAGTPGRASGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
121 PVGAVPIRSSPARSAPATRAATRESPTGSLPKFTWRBQKQPLIGCVLLIALVSLIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTGSLPKFTWRBQKQPLIGCVLLIALVSLIL 180
181 FQFQGHGTGIRYKQKRESCEPHAVRCGVNDCKLKSDELGVCRDMDKSLIKYSGSSHQ 240
181 FQFQGHGTGIRYKQKRESCEPHAVRCGVNDCKLKSDELGVCRDMDKSLIKYSGSSHQ 240
241 WLPIKSSNMNDVSEKTCQOLGFEASRHTTEVAHRDFANSFSILRYNSTIOESLHRSCEP 300
```

Db 241 WLPICSSNMNDYSSEKTCQOLGFESARHTEVAHRDFANSFSLRYNSTIOESLHRSBCP 300
Qy 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAA 360
Db 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAA 360
Qy 361 HCFVTRREKVLKGMKVYAGTNSLHQLPEAASIAEIIINSNTDEEDYDIALMLRSKPLT 420
Db 361 HCFVTRREKVLKGMKVYAGTNSLHQLPEAASIAEIIINSNTDEEDYDIALMLRSKPLT 420
Qy 421 LSAIHHPACLPMHQOTFSINETCWTGFGKTRTDDKTSPLREOVNVLIDFKKCNLDLV 480
Db 421 LSAIHHPACLPMHQOTFSINETCWTGFGKTRTDDKTSPLREOVNVLIDFKKCNLDLV 480
Qy 481 YDSYLTFRMCAGLRGGRDSCQDSDGGLVCEONNRWYLAGVTSWGTGCGQRKPKGYT 540
Db 481 YDSYLTFRMCAGLRGGRDSCQDSDGGLVCEONNRWYLAGVTSWGTGCGQRKPKGYT 540
Qy 541 KTEVLPWISKMESEVFRKS 562
Db 541 KTEVLPWISKMESEVFRKS 562

RESULT 2

US-10-806-370-12
Sequence 12, Application US/10806370
Publication No. US20040209327A1
GENERAL INFORMATION:
APPLICANT: Gedrich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/10/806.370
CURRENT FILING DATE: 2004-03-23
PRIOR APPLICATION NUMBER: US/09/879.792
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT
PRIOR FILING DATE: 2001-06-12 (Docket No. LIO-81-WO)
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-10-806-370-12

Query Match 100.0%; Score 2999; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 1 MERDSHGNSAPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
Qy 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 120
Db 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 120
Qy 121 PVGAVPIRSSPARAPATRAATRESPTSLPKFTWREGKQPLIGCVLLILALVVSLLITL 180
Db 121 PVGAVPIRSSPARAPATRAATRESPTSLPKFTWREGKQPLIGCVLLILALVVSLLITL 180
Qy 181 FQFMQGHGTGIYKQORSCPKHAIVRCGVVDCXKSDGLGCVRFDMKSLIKTISGSSHQ 240
Db 181 FQFMQGHGTGIYKQORSCPKHAIVRCGVVDCXKSDGLGCVRFDMKSLIKTISGSSHQ 240

Qy 241 WLPICSSNMNDYSSEKTCQOLGFESARHTEVAHRDFANSFSLRYNSTIOESLHRSBCP 300
Db 241 WLPICSSNMNDYSSEKTCQOLGFESARHTEVAHRDFANSFSLRYNSTIOESLHRSBCP 300
Qy 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAA 360
Db 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAA 360
Qy 361 HCFVTRREKVLKGMKVYAGTNSLHQLPEAASIAEIIINSNTDEEDYDIALMLRSKPLT 420
Db 361 HCFVTRREKVLKGMKVYAGTNSLHQLPEAASIAEIIINSNTDEEDYDIALMLRSKPLT 420
Qy 421 LSAIHHPACLPMHQOTFSINETCWTGFGKTRTDDKTSPLREOVNVLIDFKKCNLDLV 480
Db 421 LSAIHHPACLPMHQOTFSINETCWTGFGKTRTDDKTSPLREOVNVLIDFKKCNLDLV 480
Qy 481 YDSYLTFRMCAGLRGGRDSCQDSDGGLVCEONNRWYLAGVTSWGTGCGQRKPKGYT 540
Db 481 YDSYLTFRMCAGLRGGRDSCQDSDGGLVCEONNRWYLAGVTSWGTGCGQRKPKGYT 540
Qy 541 KTEVLPWISKMESEVFRKS 562
Db 541 KTEVLPWISKMESEVFRKS 562

RESULT 3

US-10-428-275-130
Sequence 130, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428.275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 130
LENGTH: 569
TYPE: PRT
ORGANISM: Homo sapiens
US-10-428-275-130

Query Match 100.0%; Score 2999; DB 4; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 5 MERDSHGNSAPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 64
Qy 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 120
Db 65 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 124
Qy 121 PVGAVPIRSSPARAPATRAATRESPTSLPKFTWREGKQPLIGCVLLILALVVSLLITL 180

Db 125 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLIL 184
Qy 181 FQWOGHTGIRYKQORSCPKHAIVCDGVNDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 185 FQWOGHTGIRYKQORSCPKHAIVCDGVNDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 244
Qy 241 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Db 245 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 304
Qy 301 SQRYSIQCHSGCGRAMTGRIVGALASDSKMPQVSLHFQTHIICGTLIDAOQWVLTAA 360
Db 305 SQRYSIQCHSGCGRAMTGRIVGALASDSKMPQVSLHFQTHIICGTLIDAOQWVLTAA 364
Qy 361 HCFVTRKEKULEGKVVAGTNSLHQLPEASIAEIIINSNTDEDDYDIALMLSKPLT 420
Db 365 HCFVTRKEKULEGKVVAGTNSLHQLPEASIAEIIINSNTDEDDYDIALMLSKPLT 424
Qy 421 LSAHIHPACLPMHGQTSLSNETCMTGFGKTRTDDKTSPLREYQVNLIDFKKNDYLV 480
Db 425 LSAHIHPACLPMHGQTSLSNETCMTGFGKTRTDDKTSPLREYQVNLIDFKKNDYLV 484
Qy 481 YDSYLTFRMPCAGDLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
Db 485 YDSYLTFRMPCAGDLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 544
Qy 541 KTEVULPMIYSKMESEVFRKS 562
Db 545 KTEVULPMIYSKMESEVFRKS 566

RESULT 4

US-10-156-214A-24
; Sequence 24, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempile
; APPLICANT: George P. Vlasuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Siev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156,214A
; CURRENT FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 562
; TYPE: PRT
; ORGANISM: homo sapien
; FEATURE:
; NAME/KEY: protease domain of endochelase 2
; LOCATION: (321)..(562)
US-10-156-214A-24

Query Match 99.7%; Score 2991; DB 4; Length 562;
Best Local Similarity 99.8%; Pred. No. 7,2e-178;
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MERDShGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
Db 1 MERDShGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
Qy 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Db 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Qy 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLIL 180
Db 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLIL 180

Qy 181 FQWOGHTGIRYKQORSCPKHAIVCDGVNDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 181 FQWOGHTGIRYKQORSCPKHAIVCDGVNDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Qy 241 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Qy 301 SQRYSIQCHSGCGRAMTGRIVGALASDSKMPQVSLHFQTHIICGTLIDAOQWVLTAA 360
Db 301 SQRYSIQCHSGCGRAMTGRIVGALASDSKMPQVSLHFQTHIICGTLIDAOQWVLTAA 360
Qy 361 HCFVTRKEKULEGKVVAGTNSLHQLPEASIAEIIINSNTDEDDYDIALMLSKPLT 420
Db 361 HCFVTRKEKULEGKVVAGTNSLHQLPEASIAEIIINSNTDEDDYDIALMLSKPLT 420
Qy 421 LSAHIHPACLPMHGQTSLSNETCMTGFGKTRTDDKTSPLREYQVNLIDFKKNDYLV 480
Db 421 LSAHIHPACLPMHGQTSLSNETCMTGFGKTRTDDKTSPLREYQVNLIDFKKNDYLV 480
Qy 481 YDSYLTFRMPCAGDLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
Db 481 YDSYLTFRMPCAGDLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
Qy 541 KTEVULPMIYSKMESEVFRKS 562
Db 541 KTEVULPMIYSKMESEVFRKS 562

RESULT 5

US-10-916-758-94
; Sequence 94, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELIAL-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-916-758-94

Query Match 99.7%; Score 2991; DB 5; Length 562;
Best Local Similarity 99.8%; Pred. No. 7,2e-178;
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MERDShGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
Db 1 MERDShGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
Qy 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Db 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Qy 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLIL 180
Db 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLIL 180
Qy 181 FQWOGHTGIRYKQORSCPKHAIVCDGVNDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 181 FQWOGHTGIRYKQORSCPKHAIVCDGVNDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240

QY 241 WLPICSSNMNDYSSEKTCOOLGFESAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSBCP 300
Db 241 WLPICSSNMNDYSSEKTCOOLGFESAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSBCP 300
QY 301 SORVLSLOCHSGHGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOQVLTAA 360
Db 301 SORVLSLOCHSGHGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOQVLTAA 360
QY 361 HCFVTRREKVLGKMWVYAGTSLHQLPEAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
Db 361 HCFVTRREKVLGKMWVYAGTSLHQLPEAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
QY 421 LSAHHPACLPMHQGTSLNETCWTITGFKTRETDDKTSPLREYQVNLIDFKCNDLV 480
Db 421 LSAHHPACLPMHQGTSLNETCWTITGFKTRETDDKTSPLREYQVNLIDFKCNDLV 480
QY 481 YDSYLTFRMCAGLRGRDSCQDSGGPLVCEQNNRYLAGVTSWGTGCGQRNKPgyt 540
Db 481 YDSYLTFRMCAGLRGRDSCQDSGGPLVCEQNNRYLAGVTSWGTGCGQRNKPgyt 540
QY 541 KTEVLPMIYSKMESEVRFRKS 562
Db 541 KTEVLPMIYSKMESEVRFRKS 562

RESULT 6

US-10-156-214A-26
; Sequence 26, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempile
; APPLICANT: George P. Vlaeuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Slev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; TITLE OF INVENTION: Theetof
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156,214A
; CURRENT FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: protease domain
; LOCATION: (321)..(688)
US-10-156-214A-26

Query Match 98.8%; Score 2963; DB 4; Length 688;
Best Local Similarity 99.3%; Pred. No. 4.9e-176;
Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNSAPARTPSAGASPAQSPAGTPPGRASPAQASPAQSPAGTPPGRASPAQAS 60
Db 1 MERDSHGNSAPARTPSAGASPAQSPAGTPPGRASPAQASPAQSPAGTPPGRASPAQAS 60
QY 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYVLYRAT 120
Db 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYVLYRAT 120
QY 121 PVGAVPIRSSPARAPATRAITRESPTSLPKFTWRBQOKPLIGCVLLIALVSLIIL 180
Db 121 PVGAVPIRSSPARAPATRAITRESPTSLPKFTWRBQOKPLIGCVLLIALVSLIIL 180
QY 181 FQFQGHGTGIRYKQRESCKHAIVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 181 FQFQGHGTGIRYKQRESCKHAIVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
QY 241 WLPICSSNMNDYSSEKTCOOLGFESAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSBCP 300

Db 241 WLPICSSNMNDYSSEKTCOOLGFESAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSBCP 300
QY 301 SORVLSLOCHSGHGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOQVLTAA 360
Db 301 SORVLSLOCHSGHGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOQVLTAA 360
QY 361 HCFVTRREKVLGKMWVYAGTSLHQLPEAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
Db 361 HCFVTRREKVLGKMWVYAGTSLHQLPEAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
QY 421 LSAHHPACLPMHQGTSLNETCWTITGFKTRETDDKTSPLREYQVNLIDFKCNDLV 480
Db 421 LSAHHPACLPMHQGTSLNETCWTITGFKTRETDDKTSPLREYQVNLIDFKCNDLV 480
QY 481 YDSYLTFRMCAGLRGRDSCQDSGGPLVCEQNNRYLAGVTSWGTGCGQRNKPgyt 540
Db 481 YDSYLTFRMCAGLRGRDSCQDSGGPLVCEQNNRYLAGVTSWGTGCGQRNKPgyt 540
QY 541 KTEVLPMIYSKMESEVR 558
Db 541 KTEVLPMIYSKMESEVR 558

RESULT 7

US-10-916-758-2
; Sequence 2, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELIAL-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-916-758-2

Query Match 98.8%; Score 2963; DB 5; Length 688;
Best Local Similarity 99.3%; Pred. No. 4.9e-176;
Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNSAPARTPSAGASPAQSPAGTPPGRASPAQASPAQSPAGTPPGRASPAQAS 60
Db 1 MERDSHGNSAPARTPSAGASPAQSPAGTPPGRASPAQASPAQSPAGTPPGRASPAQAS 60
QY 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYVLYRAT 120
Db 61 PAGTPPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYVLYRAT 120
QY 121 PVGAVPIRSSPARAPATRAITRESPTSLPKFTWRBQOKPLIGCVLLIALVSLIIL 180
Db 121 PVGAVPIRSSPARAPATRAITRESPTSLPKFTWRBQOKPLIGCVLLIALVSLIIL 180
QY 181 FQFQGHGTGIRYKQRESCKHAIVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 181 FQFQGHGTGIRYKQRESCKHAIVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
QY 241 WLPICSSNMNDYSSEKTCOOLGFESAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSBCP 300
Db 241 WLPICSSNMNDYSSEKTCOOLGFESAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSBCP 300
QY 301 SORVLSLOCHSGHGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOQVLTAA 360

Db 301 SQRYSLQCSHCHGRAMTGRIVGGLASDSKMPQVSLHFGTTHICGGLIDAQWVLTAA 360
 QY HCFVTRKVLGKMKVAGTGNLHQLPEASIAEIIINSNTDEEDDYDIALMRSLKPLT 420
 Db 361 HCFVTRKVLGKMKVAGTGNLHQLPEASIAEIIINSNTDEEDDYDIALMRSLKPLT 420
 QY 421 LSAIHHPACLPFHGOTFSLNETCMTITGFKTRTDDKTSPLREYQVNLIDPKKNDLV 480
 Db 421 LSAIHHPACLPFHGOTFSLNETCMTITGFKTRTDDKTSPLREYQVNLIDPKKNDLV 480
 QY 481 YDSYLTPEPMACGLDGRGDSGCGSGPLVCEONNRMYLAGVTSWGTCCGQRNKPYYT 540
 Db 481 YDSYLTPEPMACGLDGRGDSGCGSGPLVCEONNRMYLAGVTSWGTCCGQRNKPYYT 540
 QY 541 KYTEVLPWYSKMESEVR 558
 Db 541 KYTEVLPWYSKMESEVR 558

RESULT 8

US-10-353-690-100
 ; Sequence 100, Application US/10353690
 ; Publication No. US20030215840A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Logan, Thomas Joseph
 ; APPLICANT: Chun, Miyoung
 ; APPLICANT: Galvin, Katherine M.
 ; APPLICANT: Healy, Aileen
 ; APPLICANT: Acton, Susan L.
 ; APPLICANT: Donoghue, Mary
 ; APPLICANT: Stagliano, Nancy
 ; APPLICANT: Perodini, Jacqueline
 ; APPLICANT: Rodigue-Way, Amelie
 ; TITLE OF INVENTION: Methods and compositions for treating
 ; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
 ; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
 ; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
 ; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
 ; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
 ; TITLE OF INVENTION: 283, 2554, 9464, 17799, 2686, 43848, 32135, 12208, 2914,
 ; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
 ; FILE REFERENCE: MP102-018P1RMONIM
 ; FILE REFERENCE: 9792, 15400, 1452 or 6585 molecules
 ; CURRENT APPLICATION NUMBER: US/10/353,690
 ; CURRENT FILING DATE: 2003-01-29
 ; PRIOR APPLICATION NUMBER: 60/353,224
 ; PRIOR FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: 60/364,529
 ; PRIOR FILING DATE: 2002-03-15
 ; PRIOR APPLICATION NUMBER: 60/373,861
 ; PRIOR FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: 60/376,287
 ; PRIOR FILING DATE: 2002-04-29
 ; PRIOR APPLICATION NUMBER: 60/386,080
 ; PRIOR FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: 60/390,971
 ; PRIOR FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: 60/394,130
 ; PRIOR FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: 60/394,797
 ; PRIOR FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: 60/404,904
 ; PRIOR FILING DATE: 2002-08-21
 ; PRIOR APPLICATION NUMBER: 60/405,450
 ; PRIOR FILING DATE: 2002-08-23
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 100
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-353-690-100

Query Match 98.6%; Score 2956; DB 4; Length 581;
 Best Local Similarity 99.8%; Pred. No. 1,1e-175;
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERDSHGASPARTPSPAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
 Db 1 MERDSHGASPARTPSPAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
 QY 61 PACTPPGRASPGRASPAQASPARASPALASISRSSSSGSSSARASVTSSTRVLYVAT 120
 Db 61 PACTPPGRASPGRASPAQASPARASPALASISRSSSSGSSSARASVTSSTRVLYVAT 120
 QY 121 PVGAVPIRSSPARASPARATPATRESPTSLPKFTREGOKOPLIGCVLLALVLSLTL 180
 Db 121 PVGAVPIRSSPARASPARATPATRESPTSLPKFTREGOKOPLIGCVLLALVLSLTL 180
 QY 181 FQFWGHTGIRYKQRESCEPHAVRCGVVDCXLSDELGCVRPDMKSLIKIYSSSHQ 240
 Db 181 FQFWGHTGIRYKQRESCEPHAVRCGVVDCXLSDELGCVRPDMKSLIKIYSSSHQ 240
 QY 241 WLPICSSNMWDSYSEKTCQOQGFSSAHTTVAARDPANSFSLRYNSTIOESLHRSCEP 300
 Db 241 WLPICSSNMWDSYSEKTCQOQGFSSAHTTVAARDPANSFSLRYNSTIOESLHRSCEP 300
 QY 301 SQRYSLQCSHCHGRAMTGRIVGGLASDSKMPQVSLHFGTTHICGGLIDAQWVLTAA 360
 Db 301 SQRYSLQCSHCHGRAMTGRIVGGLASDSKMPQVSLHFGTTHICGGLIDAQWVLTAA 360
 QY 361 HCFVTRKVLGKMKVAGTGNLHQLPEASIAEIIINSNTDEEDDYDIALMRSLKPLT 420
 Db 361 HCFVTRKVLGKMKVAGTGNLHQLPEASIAEIIINSNTDEEDDYDIALMRSLKPLT 420
 QY 421 LSAIHHPACLPFHGOTFSLNETCMTITGFKTRTDDKTSPLREYQVNLIDPKKNDLV 480
 Db 421 LSAIHHPACLPFHGOTFSLNETCMTITGFKTRTDDKTSPLREYQVNLIDPKKNDLV 480
 QY 481 YDSYLTPEPMACGLDGRGDSGCGSGPLVCEONNRMYLAGVTSWGTCCGQRNKPYYT 540
 Db 481 YDSYLTPEPMACGLDGRGDSGCGSGPLVCEONNRMYLAGVTSWGTCCGQRNKPYYT 540
 QY 541 KYTEVLPWYSKME 554
 Db 541 KYTEVLPWYSKME 554

RESULT 9

US-10-428-275-104
 ; Sequence 104, Application US/10428275
 ; Publication No. US20040067505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alvarez et al.
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
 ; FILE REFERENCE: 21402-585
 ; CURRENT APPLICATION NUMBER: US/10/428,275
 ; CURRENT FILING DATE: 2003-05-01
 ; PRIOR APPLICATION NUMBER: 09/966545
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 09/544511
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: 60/128514
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/569269
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIOR APPLICATION NUMBER: 60/134315
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: 09/619252
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/789390
 ; PRIOR FILING DATE: 2001-02-23
 ; PRIOR APPLICATION NUMBER: 60/185548
 ; PRIOR FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 450

SOFTWARE: Curaseqblast version 0.1
 SEQ ID NO 104
 LENGTH: 586
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-428-275-104

Query Match 98.3%; Score 2948.5; DB 4; Length 586;
 Best Local Similarity 99.1%; Pred. No. 3.3e-175;
 Matches 554; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

QY 1 MERDSHGNAAPRTSPASAPPAQSPAGTPPGRASPAQSPAPQASPAQSPAGTPPGRASPAQAS 60
DB 1 MERDSHGNAAPRTSPASAPPAQSPAGTPPGRASPAQSPAPQASPAQSPAGTPPGRASPAQAS 60
QY 61 PAGTPPGRASPGR-----ASPAQASPARASPALASLSRSSSGRSSSARSASAVTTSPTRY 115
DB 61 PAGTPPGRASPGRASPAPQASPAQASPARASPALASLSRSSSGRSSSARSASAVTTSPTRY 120
QY 116 LVRATPVGAVPFIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 175
DB 121 LVRATPVGAVPFIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 180
QY 176 SLIIIFQWOGHTGIRYKQRESCEPKAVRCDGVYDCKLKSDCLGCVAFMDKSLIKIYS 235
DB 181 SLIIIFQWOGHTGIRYKQRESCEPKAVRCDGVYDCKLKSDCLGCVAFMDKSLIKIYS 240
QY 236 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 295
DB 241 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 300
QY 296 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTHICGTLIDAQW 355
DB 301 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTHICGTLIDAQW 360
QY 361 VLTAAHCFVTRREKYLEGMKYAGTSLNHLQPEASIAEIIINSYTTBEDDYDIALMRL 420
DB 366 VLTAAHCFVTRREKYLEGMKYAGTSLNHLQPEASIAEIIINSYTTBEDDYDIALMRL 415
QY 416 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDFKKC 475
DB 421 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDFKKC 480
QY 476 NDLYVDSYLTFRMWCAGDLRGGRSCQDGGGPLVCEQNNRWYLAGVTSMTGCGGRNK 535
DB 481 NDLYVDSYLTFRMWCAGDLRGGRSCQDGGGPLVCEQNNRWYLAGVTSMTGCGGRNK 540
QY 536 PGVYTKVTEVLPWYISKME 554
DB 541 PGVYTKVTEVLPWYISKME 559

```

RESULT 10
 US-10-428-275-124
 Sequence 124, Application US/10428275
 Publication No. US20040067505A1
 GENERAL INFORMATION:
 APPLICANT: Alvarez et al.
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
 FILE REFERENCE: 21402-585
 CURRENT APPLICATION NUMBER: US/10/428,275
 CURRENT FILING DATE: 2003-05-01
 PRIOR APPLICATION NUMBER: 09/566545
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 09/544511
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/128514
 PRIOR FILING DATE: 1999-04-09
 PRIOR APPLICATION NUMBER: 09/569269
 PRIOR FILING DATE: 2000-05-11
 PRIOR APPLICATION NUMBER: 60/134315
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 09/619252

PRIOR FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 09/789390
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/185548
 PRIOR FILING DATE: 2000-02-25
 NUMBER OF SEQ ID NOS: 450
 SOFTWARE: Curaseqblast version 0.1
 SEQ ID NO 124
 LENGTH: 586
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-428-275-124

Query Match 98.3%; Score 2948.5; DB 4; Length 586;
 Best Local Similarity 99.1%; Pred. No. 3.3e-175;
 Matches 554; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

QY 1 MERDSHGNAAPRTSPASAPPAQSPAGTPPGRASPAQSPAPQASPAQSPAGTPPGRASPAQAS 60
DB 1 MERDSHGNAAPRTSPASAPPAQSPAGTPPGRASPAQSPAPQASPAQSPAGTPPGRASPAQAS 60
QY 61 PAGTPPGRASPGR-----ASPAQASPARASPALASLSRSSSGRSSSARSASAVTTSPTRY 115
DB 61 PAGTPPGRASPGRASPAQASPAQASPARASPALASLSRSSSGRSSSARSASAVTTSPTRY 120
QY 116 LVRATPVGAVPFIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 175
DB 121 LVRATPVGAVPFIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 180
QY 176 SLIIIFQWOGHTGIRYKQRESCEPKAVRCDGVYDCKLKSDCLGCVAFMDKSLIKIYS 235
DB 181 SLIIIFQWOGHTGIRYKQRESCEPKAVRCDGVYDCKLKSDCLGCVAFMDKSLIKIYS 240
QY 236 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 295
DB 241 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 300
QY 296 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTHICGTLIDAQW 355
DB 301 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTHICGTLIDAQW 360
QY 361 VLTAAHCFVTRREKYLEGMKYAGTSLNHLQPEASIAEIIINSYTTBEDDYDIALMRL 420
DB 366 VLTAAHCFVTRREKYLEGMKYAGTSLNHLQPEASIAEIIINSYTTBEDDYDIALMRL 415
QY 416 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDFKKC 475
DB 421 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDFKKC 480
QY 476 NDLYVDSYLTFRMWCAGDLRGGRSCQDGGGPLVCEQNNRWYLAGVTSMTGCGGRNK 535
DB 481 NDLYVDSYLTFRMWCAGDLRGGRSCQDGGGPLVCEQNNRWYLAGVTSMTGCGGRNK 540
QY 536 PGVYTKVTEVLPWYISKME 554
DB 541 PGVYTKVTEVLPWYISKME 559

```

RESULT 11
 US-10-428-275-120
 Sequence 120, Application US/10428275
 Publication No. US20040067505A1
 GENERAL INFORMATION:
 APPLICANT: Alvarez et al.
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
 FILE REFERENCE: 21402-585
 CURRENT APPLICATION NUMBER: US/10/428,275
 CURRENT FILING DATE: 2003-05-01
 PRIOR APPLICATION NUMBER: 09/566545
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 09/544511
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/128514

| Query Match | Similarity | 98.1% | Score | 2943.5 | DB 4 | Length | 586 |
|-------------|------------|------------------------------|-----------------------------|------------|-----------------|------------------|----------------------|
| Best Local | Similarity | 98.9% | Pred. No. | 6.8e-175 | | | |
| Matches | 553 | Conservative | 0 | Mismatches | 1 | Indels | 5 |
| | | | | | | | Gaps |
| | | | | | | | 1 |
| QY | 1 | MERDSHGNASPARFPYSPAGASPAQAS | PAGTPPGRASPAQAS | EPAQAS | PAGTPPGRASPAQAS | 60 | |
| DB | 1 | MERSHGNASPARFPSPAGASPAQAS | PAGTPPGRASPAQAS | EPAQAS | PAGTPPGRASPAQAS | 60 | |
| QY | 61 | PAGTPPGRASFGF---- | ASPAQAS | SPAPAS | PALASLS | BSRSSGSSSSARASAS | VTTSPPRY 115 |
| DB | 61 | PAGTPPGRASFGSPASPAQAS | SPAQAS | SPARAS | PALASLS | BSRSSGSSSSARASAS | VTTSPPRY 120 |
| QY | 116 | LVRATPVGAVYIRASPARSAPATATPES | PETSLPKFTWRGQOLPLIGCVLLIALV | 175 | | | |
| DB | 121 | LVRATPVGAVYIRSSPARSAPATATPES | PETSLPKFTWRGQOLPLIGCVLLIALV | 180 | | | |
| QY | 176 | SLILFQFQWQHTGIRYKEQRES | CPKIAVRC | CDGVVDC | CKLSDEL | GCYAF | FDMDKSLTKYS 235 |
| DB | 181 | SLILFQFQWQHTGIRYKEQRES | CPKIAVRC | CDGVVDC | CKLSDEL | GCYAF | FDMDKSLTKYS 240 |
| QY | 236 | GSSHOMLPICSSNMNDYSSEKTC | COOLGFES | AKHTTEVA | HRPANS | FSILTR | NSTQESLH 295 |
| DB | 241 | GSSHOMLPICSSNMNDYSSEKTC | COOLGFES | AKHTTEVA | HRPANS | FSILTR | NSTQESLH 300 |
| QY | 296 | RSECSQRYILQCSHCGLRAMTGR | IYOGALAS | DSKMPWQVSL | HFGTTH | ICGTL | ILDNQW 355 |
| DB | 301 | RSECSQRYILQCSHCGLRAMTGR | IYOGALAS | DSKMPWQVSL | HFGTTH | ICGTL | ILDNQW 360 |
| QY | 356 | VLTAAHCFVTVREKYLEGKMKV | YAGTSLH | LOLPEAAS | IAEII | INSNTY | DEEDDYDIALMRL 415 |
| DB | 361 | VLTAAHCFVTVREKYLEGKMKV | YAGTSLH | LOLPEAAS | IAEII | INSNTY | DEEDDYDIALMRL 420 |
| QY | 416 | SKPLTSLSHIHIPACLPNHGOT | FSINETC | WITIPG | KTRTDD | TKTSP | PLREVOVNLIDPFKKC 475 |
| DB | 421 | SKPLTSLSHIHIPACLPNHGOT | FSINETC | WITIPG | KTRTDD | TKTSP | PLREVOVNLIDPFKKC 480 |
| QY | 476 | NDLYLVYDSYILPRMWCAG | DLRGGRDS | CCGDSG | GPLVCE | ONNRWY | LAGVTS |
| DB | 481 | NDLYLVYDSYILPRMWCAG | DLRGGRDS | CCGDSG | GPLVCE | ONNRWY | LAGVTS |
| QY | 536 | PGVYTKVTETVLPWYSKME | 554 | | | | |
| DB | 541 | PGVYTKVTETVLPWYSKME | 559 | | | | |

```
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT FILING DATE: 2003-05-01
CURRENT APPLICATION NUMBER: US/10/428,275
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Curaseqlist version 0.1
SEQ ID NO 140
LENGTH: 586
TYPE: PRT
ORGANISM: Homo sapiens
US-10-428-275-140
```

```
Query Match 98.1%; Score 2942.5; DB 4; Length 586;
Best Local Similarity 98.9%; Pred. No. 7.8e-175;
Matches 553; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
DB 1 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
QY 61 PAGTPPGASPGR-----ASPAQASPARASPALASLSRSSSGRSSASASAVTTSPTVY 115
DB 61 PAGTPPGASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
QY 116 LVRATPVGAVPIRSSPARASAPATRAATRESPTSLPFTWRBQKQLPLIGCVLLIALV 175
DB 121 LVRATPVGAVPIRSSPARASAPATRAATRESPTSLPFTWRBQKQLPLIGCVLLIALV 180
QY 176 SLILIFQWQHTGIRYKEQRESCPKHAIVRCGVVDCXKLSDELGCVRFDMDKSLKITS 235
DB 181 SLILIFQWQHTGIRYKEQRESCPKHAIVRCGVVDCXKLSDELGCVRFDMDKSLKITS 240
QY 236 GSSHOMLPTICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIQESLH 295
DB 241 GSSHOMLPTICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIQESLH 300
QY 296 RSECPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAQW 355
DB 301 RSECPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAQW 360
QY 356 VLTAAHCFVTRREKYLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDIAMRL 415
DB 361 VLTAAHCFVTRREKYLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDIAMRL 420
QY 416 SKPLTSLAHIPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREVOVNLIDFKKC 475
DB 421 SKPLTSLAHIPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREVOVNLIDFKKC 480
QY 476 NDLYVYDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEONNRWYLAGVTSWGTGCGQRNK 535
DB 481 NDLYVYDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEONNRWYLAGVTSWGTGCGQRNK 540
QY 536 PGVYTKVTEVLPWYSKME 554
DB 540 PGVYTKVTEVLPWYSKME 554
```

```
DB 541 PGVYTKVTEVLPWYSKME 559
RESULT 14
US-09-888-615-104
Sequence 104, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: MAYHE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 104
LENGTH: 537
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-104
```

```
Query Match 92.5%; Score 2775.5; DB 3; Length 537;
Best Local Similarity 92.0%; Pred. No. 1.8e-164;
Matches 526; Conservative 0; Mismatches 1; Indels 45; Gaps 2;

QY 1 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
DB 1 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
QY 61 PAGTPPGASPGR-----ASPAQASPARASPALASLSRSSSGRSSASARASVTS 110
DB 61 PAGTPPGASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
QY 111 PTRYIVRATPVGAVPIRSSPARASAPATRAATRESPTSLPFTWRBQKQLPLIGCVLL 170
DB 121 PTRYIVRATPVGAVPIRSSPARASAPATRAATRESPTSLPFTWRBQKQLPLIGCVLL 155
QY 171 IALVSLIILFQWQHTGIRYKEQRESCPKHAIVRCGVVDCXKLSDELGCVRFDMDKSL 230
DB 156 IALVSLIILFQWQHTGIRYKEQRESCPKHAIVRCGVVDCXKLSDELGCVRFDMDKSL 205
QY 231 LKIYSGSSHOMLPTICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTI 290
DB 206 LKIYSGSSHOMLPTICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTI 265
QY 291 QESLHRSRCPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGL 350
DB 266 QESLHRSRCPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGL 325
QY 351 IDAQWVLTAAHCFVTRREKYLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDI 410
DB 326 IDAQWVLTAAHCFVTRREKYLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDI 385
QY 411 ALMLRSLKPLTSLAHIPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREVOVNL 470
DB 386 ALMLRSLKPLTSLAHIPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREVOVNL 445
QY 471 DFKKCNQDLYVYDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEONNRWYLAGVTSWGTGC 530
DB 446 DFKKCNQDLYVYDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEONNRWYLAGVTSWGTGC 505
QY 531 GQRNKPQVYTKVTEVLPWYSKMESEVFRKS 562
DB 506 GQRNKPQVYTKVTEVLPWYSKMESEVFRKS 537
```


RESULT 15

US-10-428-275-128
; Sequence 128, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428,275
; PRIOR FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 128
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-275-128

Query Match 92.5%; Score 2775.5; DB 4; Length 537;
Best Local Similarity 92.0%; Pred. No. 1.8e-164;
Matches 526; Conservative 0; Mismatches 1; Indels 45; Gaps 2;

QY 1 MERDSHGASPARTSPAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
DB 1 MERDSHGASPARTSPAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
QY 61 PAGTPGRASPCR-----ASPAQASPARASPALASLSRSSSSGRSSSARSASVTTTS 110
DB 61 PAGTPGRASPCRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
QY 111 PTRVYLVPATPVGAIVRIRSSPARSAPATRTATRESPTSLPKFTWREGQQLPLICVLLL 170
DB 121 PTRVYLVPATPVGAIVRIRSSPARSAPATRTATRESPTSLPKFTWREGQQLPLICVLLL 155
QY 171 IALVSLILFQFMOGHTGIRYKEQRESCPKHVRCDGVNDCKLSDELGCVRFPMDKSL 230
DB 156 -----VQFMOGHTGIRYKEQRESCPKHVRCDGVNDCKLSDELGCVRFPMDKSL 205
QY 231 LKIYSGSSHQWLPICSSNMNDYSSEKTCQQLGFBSAHRTEVAHRDFANSFSLRYNSTI 290
DB 206 LKIYSGSSHQWLPICSSNMNDYSSEKTCQQLGFBSAHRTEVAHRDFANSFSLRYNSTI 265
QY 291 QESLHRSRCPSSORYISLQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGTL 350
DB 266 QESLHRSRCPSSORYISLQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGTL 325
QY 351 IDAQVNLTAHCFVTRREKVLGEMKVYAGTSMHLQPEAASIAEIIINSNYTDEEDDYDI 410
DB 326 IDAQVNLTAHCFVTRREKVLGEMKVYAGTSMHLQPEAASIAEIIINSNYTDEEDDYDI 385
QY 411 ALMRSLKPLTSAHIHPACTLPMHGOTFSLNETCMTTGREGTETDDKTSPLREVOVNL 470
DB 386 ALMRSLKPLTSAHIHPACTLPMHGOTFSLNETCMTTGREGTETDDKTSPLREVOVNL 445
QY 471 DFKKNDVLYVDSYLTPRMCAAGDLRGGRDSCGDSGGPLVCEQNNRWYLAGVTSWGTGC 530
DB 446 DFKKNDVLYVDSYLTPRMCAAGDLRGGRDSCGDSGGPLVCEQNNRWYLAGVTSWGTGC 505

QY 531 GORNKPQVYTKVTEVLPWISKMESEVFRKS 562
DB 506 GORNKPQVYTKVTEVLPWISKMESEVFRKS 537

Search completed: September 16, 2006, 03:29:18
Job time : 182 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:26:26 ; Search time 35 Seconds

(without alignments)
1134.395 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999
Sequence: 1 MERSHGNASAPRTSPSAGS.....TEVLPWIKMESEVRPRKS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 256596 seqs, 70647373 residues

Total number of hits satisfying chosen parameters: 256596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

1: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/PCY_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|-----|--------------------|
| 1 | 729.5 | 24.3 | 453 | US-10-196-749-64 |
| 2 | 726 | 24.2 | 454 | US-11-400-825-2 |
| 3 | 694 | 23.1 | 432 | US-10-196-749-310 |
| 4 | 694 | 23.1 | 432 | US-11-101-316-112 |
| 5 | 608.5 | 20.3 | 416 | US-11-376-673-112 |
| 6 | 608.5 | 20.3 | 416 | US-11-208-257-8 |
| 7 | 543.5 | 18.1 | 355 | US-11-293-697-3002 |
| 8 | 525.5 | 17.5 | 255 | US-11-254-185-6 |
| 9 | 525.5 | 17.5 | 255 | US-11-254-185-38 |
| 10 | 525.5 | 17.5 | 255 | US-11-253-869-6 |
| 11 | 520.5 | 17.4 | 423 | US-10-196-749-310 |
| 12 | 520.5 | 17.4 | 423 | US-10-196-749-310 |
| 13 | 520.5 | 17.4 | 423 | US-11-101-316-106 |
| 14 | 520.5 | 17.4 | 423 | US-11-376-673-106 |
| 15 | 518.5 | 17.3 | 853 | US-10-530-187-242 |
| 16 | 518.5 | 17.3 | 855 | US-11-254-185-2 |
| 17 | 518.5 | 17.3 | 855 | US-11-253-869-2 |
| 18 | 518.5 | 17.3 | 855 | US-11-303-608-1 |
| 19 | 515.5 | 17.2 | 311 | US-11-327-490-41 |
| 20 | 514 | 17.1 | 812 | US-11-318-939-7 |
| 21 | 510.5 | 17.0 | 343 | US-11-359-534-3 |
| 22 | 506 | 16.9 | 338 | US-11-318-939-10 |
| 23 | 506 | 16.9 | 531 | US-10-522-668-2 |
| 24 | 505 | 16.8 | 276 | US-11-327-490-35 |
| 25 | 498.5 | 16.6 | 339 | US-11-404-745-2 |

| | | | | | |
|----|-------|------|-----|------------------|-------------------|
| 26 | 495 | 16.5 | 812 | US-11-318-939-12 | Sequence 12, Appl |
| 27 | 493.5 | 16.5 | 790 | US-11-350-703-1 | Sequence 1, Appl |
| 28 | 493.5 | 16.5 | 791 | US-11-318-939-6 | Sequence 8, Appl |
| 29 | 493 | 16.4 | 333 | US-11-318-939-8 | Sequence 2, Appl |
| 30 | 492.5 | 16.4 | 810 | US-11-431-526-2 | Sequence 2, Appl |
| 31 | 492.5 | 16.4 | 810 | US-11-431-663-2 | Sequence 2, Appl |
| 32 | 485.5 | 16.2 | 311 | US-11-359-858-2 | Sequence 2, Appl |
| 33 | 480.5 | 16.0 | 262 | US-11-359-554-2 | Sequence 13, Appl |
| 34 | 480 | 16.0 | 790 | US-11-318-939-13 | Sequence 36, Appl |
| 35 | 479 | 16.0 | 245 | US-11-254-185-36 | Sequence 36, Appl |
| 36 | 479 | 16.0 | 245 | US-11-253-869-36 | Sequence 36, Appl |
| 37 | 479 | 16.0 | 275 | US-10-530-798-24 | Sequence 25, Appl |
| 38 | 478 | 15.9 | 275 | US-10-530-798-25 | Sequence 25, Appl |
| 39 | 476.5 | 15.9 | 272 | US-11-327-490-37 | Sequence 37, Appl |
| 40 | 475.5 | 15.9 | 272 | US-11-359-554-6 | Sequence 6, Appl |
| 41 | 473 | 15.8 | 275 | US-10-530-798-26 | Sequence 26, Appl |
| 42 | 472.5 | 15.8 | 810 | US-11-318-939-11 | Sequence 11, Appl |
| 43 | 470.5 | 15.7 | 272 | US-11-327-490-40 | Sequence 40, Appl |
| 44 | 466 | 15.5 | 655 | US-11-242-617-1 | Sequence 1, Appl |
| 45 | 465 | 15.5 | 275 | US-10-530-798-23 | Sequence 23, Appl |

ALIGNMENTS

RESULT 1
US-10-196-749-64
Sequence 64, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desroyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C340
CURRENT FILING DATE: 2002-07-16
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 64
LENGTH: 453
TYPE: PRT
ORGANISM: Homo Sapien
US-10-196-749-64

Query Match 24.3%; Score 729.5; DB 6; Length 453;
Best Local Similarity 37.8%; Pred. No. 6,8e-39;
Matches 158; Conservative 69; Mismatches 162; Indels 29; Gaps 10;

```
QY 161 LPL-----ICCVLLALVSLIILFQWOGHTGIRYKRESCPKHAYACDGVNDC 213
DB 44 LPLKEPPIIVIGIILALIGLGHFDC---SGKYRCRSSFICIELIARCDGVSDCK 99
QY 214 LKSDDELGVRFMDLSLKITYSGSSHQWLPICSSNMNDSYSEKTCQOQGFESAHRTTEVA 273
DB 100 DGEDEYRCRVGGQNAVALQVFTAA--WTKMCSDDMKGHYANVACAQLGFESYSSDNIR 157
QY 274 -----HRDFANSFILRYN--STIOBSLH-RSECPQRYISLQCSHCG-LRAMTGR1 321
DB 158 VSLRSGQFREFEVSIDHLLPDDKVTALHSHSVYRBCASGHVYTLQCTACGRRGYSRI 217
QY 322 VGGALASDSKMPQVSIHFGTTHICGGLIDAQWVLTAAHCFVTRREKVLBSMKVYAGTS 381
DB 218 VGGNNSLSQWPMQASLQFGYHLCGGSVITPLMTITAAHCYVDL--YLPSMTIQVGLV 275
QY 382 NLHOLPEAASIAE-IINSNYTDDEDDYDIALMRLSKPLTLSAHIHPACLPMHGGTFSLN 440
DB 276 SLDDNPASHLYEKVYHSHKYKPKLGDIALMKLAGPLTFNEMIQPVCLPNSBNFDPG 335
QY 441 ETCWITGFKTRBTDDKTSPLREYQVNLIDFKCNDVLYVDSYLTPLRMWCAGDLRGGRD 500
DB 336 KVCMTSGWCATRDGD-ASPLNHAVALPLISNKCINHRDVYGGIISPSMLCAGYLTGGVD 394
QY 501 SCQSGSGPLVCEQNNRWTLAGVTSWGTCCGQNRKPGYTKTEVLPWYISKMESEVR 558
DB 395 SCQSGSGPLVQGBRRLMLVGATSFGICAEVNRPGYTRVTSPLDWIHQEMERDLK 452
```

RESULT 2

US-11-400-825-2
Sequence 2, Application US/11400825
Publication No. US20060177866A1
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Underwood, Lowell J.
TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
FILE REFERENCE: D6192CIPD3
CURRENT APPLICATION NUMBER: US/11/400,825
CURRENT FILING DATE: 2006-04-07
PRIOR APPLICATION NUMBER: 10/455,720
PRIOR FILING DATE: 2003-06-05
NUMBER OF SEQ ID NOS: 154
SEQ ID NO 2
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: complete amino acid sequence of TADG-12
US-11-400-825-2

Query Match 24.2%; Score 726; DB 7; Length 454;
Best Local Similarity 37.7%; Pred. No. 1.1e-38;
Matches 154; Conservative 70; Mismatches 162; Indels 22; Gaps 9;

```
QY 164 IGVLLALVSLIILFQWOGHTGIRYKRESCPKHAYACDGVNDC LKSDDELGVCR 223
DB 55 LGIILALVSLIILFQWOGHTGIRYKRESCPKHAYACDGVNDC LKSDDELGVCR 110
QY 224 FDMWKSILKITYSGSSHQWLPICSSNMNDSYSEKTCQOQGFESAHRTTEVA-----HR 275
DB 111 VGGQNAVALQVFTAA--WTKMCSDDMKGHYANVACAQLGFESYSSDNIRVSLRSGQFRE 168
QY 276 DFANSFILRYN--STIOBSLH-RSECPQRYISLQCSHCG-LRAMTGRIVGGALASDK 331
DB 169 EFVSIIDHLLPDDKVTALHSHSVYRBCASGHVYTLQCTACGRRGYSRIVGGNMSLSL 228
```

```
QY 332 WPMQVSLHFGTTHICGGLIDAQWVLTAAHCFVTRREKVLBSMKVYAGTSNLHOLPEAAS 391
DB 229 WPMQASLQFGYHLCGGSVITPLMTITAAHCYVDL--YLPSMTIQVGLVSLDDNPASH 286
QY 392 IAE-IINSNYTDDEDDYDIALMRLSKPLTLSAHIHPACLPMHGGTFSLNETCMTITGK 450
DB 287 LVEKIVSHSKYKPKLGDIALMKLAGPLTFNEMIQPVCLPNSBNFDPGKVCMTSGMGA 346
QY 451 TRETDKTSPLREYQVNLIDFKCNDVLYVDSYLTPLRMWCAGDLRGGRDSCQSGSGPL 510
DB 347 TEDGD-ASPLNHAVALPLISNKCINHRDVYGGIISPSMLCAGYLTGGVDSQSGSGPL 405
QY 511 VCEQNNRWTLAGVTSWGTCCGQNRKPGYTKTEVLPWYISKMESEVR 558
DB 406 VQGBRRLMLVGATSFGICAEVNRPGYTRVTSPLDWIHQEMERDLK 453
```

RESULT 3

US-10-196-749-330
Sequence 330, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Auestin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 330
LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapien
US-10-196-749-330

Query Match 23.1%; Score 694; DB 6; Length 432;
Best Local Similarity 34.9%; Pred. No. 1.1e-36;
Matches 151; Conservative 73; Mismatches 161; Indels 48; Gaps 11;

```
QY 149 LPKTRRGQQLPLGCVLLIALLVSLIILFQWOGHTGIRYKRESCPKH 202
DB 149 LPKTRRGQQLPLGCVLLIALLVSLIILFQWOGHTGIRYKRESCPKH 202
```

Db 24 IPMEFR--KVGIPIIIALLSLASIIIVVLIKVIIDKYFLCGQ-----PLH 69
 Qy 203 AV-----RCDGVDDCKLSDELGC-----VRPDMDSLKIKYSSGHOMLPICSS 247
 Db 70 FIPRKQCDGELDCPLGEDEBHCYKSPFEGPAVAVRLSKDRSTLQVLDATGNMFSACFD 129
 Qy 248 NMNDYSEKTCQOOLGFESAHRTTEVAHRDPANSESIILRYNSTIOESLHRSRCPQRYISL 307
 Db 130 NFTAALAEACROMQYGS---RAVEIGPDODLDVEITENSGELMRNNSGCLSGSLVSL 186
 Qy 308 QCSHCGLRAMTGRIVGALASDSKMPWVSLHFGTTHICGTLIDAOVYTLAAHCFEYTR 367
 Db 187 HCLACGSLKTPRVVGGSEASVDSMPWQVSIQYKQHVCGSIIIDPHVLTAAHCF--RK 244
 Qy 368 EKVLEGMKVYAGTGNLHQLPEPASIAEII---NSNTDEEDVDVIALMRSLKPLTSAH 424
 Db 245 HTDVFNMKVVRAGSDKLSGSPSLA--VAKIIIEFNPWY---PKNDIDALMKQLPFTFSGT 300
 Qy 425 IHPACLPMHGQTFSLNETCWITGFKTRBDKTSPPFLREVQVNLIDPKKNDVLYVDSY 484
 Db 301 VRPCLPFPDEBLTPATPLWITGWFYKONGKSDILLQASVQYIDSTRCNADAYGGE 360
 Qy 485 LTPRMGACGLRGGRDSCQDSDGGLVCEONNRWYLAGVTSWGTGCGGRNKPQYTTKYTE 544
 Db 361 VTERMMGAGIPGEGVDTCQDSDGGLM--YQSDQHVHVGIVSGWYCGGSPSTPGYTTKVA 419
 Qy 545 VLPWYKMESEV 557
 Db 420 YLNMWYVWKAEL 432

RESULT 4

US-11-101-316-112
 ; Sequence 112, Application US/11101316
 ; Publication No. US20060099657A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
 ; TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
 ; FILE REFERENCE: P3230R1C17C1
 ; CURRENT APPLICATION NUMBER: US/11/101,316
 ; PRIOR FILING DATE: 2005-04-06
 ; PRIOR APPLICATION NUMBER: 10/063526
 ; PRIOR FILING DATE: 2002-05-03
 ; PRIOR APPLICATION NUMBER: 10/006867
 ; PRIOR FILING DATE: 2001-12-06
 ; PRIOR APPLICATION NUMBER: PCT/US00/23328
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/380137
 ; PRIOR FILING DATE: 1999-08-25
 ; PRIOR APPLICATION NUMBER: PCT/US99/12252
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 112
 ; LENGTH: 432
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-11-101-316-112

Query Match 23.1%; Score 694; DB 7; Length 432;
 Best Local Similarity 34.9%; Pred. No. 1,1e-36;
 Matches 151; Conservative 73; Mismatches 161; Indels 48; Gaps 11;
 Qy 149 LPKFTWREGQKQLPIGCVLLIALVLSLIL-----FQFWQHTGIRYKQRESRCPKH 202
 Db 24 IPMEFR--KVGIPIIIALLSLASIIIVVLIKVIIDKYFLCGQ-----PLH 69

Qy 203 AV-----RCDGVDDCKLSDELGC-----VRPDMDSLKIKYSSGHOMLPICSS 247
 Db 70 FIPRKQCDGELDCPLGEDEBHCYKSPFEGPAVAVRLSKDRSTLQVLDATGNMFSACFD 129
 Qy 248 NMNDYSEKTCQOOLGFESAHRTTEVAHRDPANSESIILRYNSTIOESLHRSRCPQRYISL 307
 Db 130 NFTAALAEACROMQYGS---RAVEIGPDODLDVEITENSGELMRNNSGCLSGSLVSL 186
 Qy 308 QCSHCGLRAMTGRIVGALASDSKMPWVSLHFGTTHICGTLIDAOVYTLAAHCFEYTR 367
 Db 187 HCLACGSLKTPRVVGGSEASVDSMPWQVSIQYKQHVCGSIIIDPHVLTAAHCF--RK 244
 Qy 368 EKVLEGMKVYAGTGNLHQLPEPASIAEII---NSNTDEEDVDVIALMRSLKPLTSAH 424
 Db 245 HTDVFNMKVVRAGSDKLSGSPSLA--VAKIIIEFNPWY---PKNDIDALMKQLPFTFSGT 300
 Qy 425 IHPACLPMHGQTFSLNETCWITGFKTRBDKTSPPFLREVQVNLIDPKKNDVLYVDSY 484
 Db 301 VRPCLPFPDEBLTPATPLWITGWFYKONGKSDILLQASVQYIDSTRCNADAYGGE 360
 Qy 485 LTPRMGACGLRGGRDSCQDSDGGLVCEONNRWYLAGVTSWGTGCGGRNKPQYTTKYTE 544
 Db 361 VTERMMGAGIPGEGVDTCQDSDGGLM--YQSDQHVHVGIVSGWYCGGSPSTPGYTTKVA 419
 Qy 545 VLPWYKMESEV 557
 Db 420 YLNMWYVWKAEL 432

RESULT 5

US-11-376-673-112
 ; Sequence 112, Application US/11376673
 ; Publication No. US20060160186A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
 ; TITLE OF INVENTION: LUNG TUMOR
 ; FILE REFERENCE: P3230R1C165C
 ; CURRENT APPLICATION NUMBER: US/11/376,673
 ; PRIOR FILING DATE: 2006-03-14
 ; PRIOR APPLICATION NUMBER: 10/063742
 ; PRIOR FILING DATE: 2002-05-09
 ; PRIOR APPLICATION NUMBER: 10/006867
 ; PRIOR FILING DATE: 2001-12-06
 ; PRIOR APPLICATION NUMBER: PCT/US00/23328
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 60/170262
 ; PRIOR FILING DATE: 1999-12-09
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 112
 ; LENGTH: 432
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-11-376-673-112

Query Match 23.1%; Score 694; DB 7; Length 432;
 Best Local Similarity 34.9%; Pred. No. 1,1e-36;
 Matches 151; Conservative 73; Mismatches 161; Indels 48; Gaps 11;
 Qy 149 LPKFTWREGQKQLPIGCVLLIALVLSLIL-----FQFWQHTGIRYKQRESRCPKH 202
 Db 24 IPMEFR--KVGIPIIIALLSLASIIIVVLIKVIIDKYFLCGQ-----PLH 69
 Qy 203 AV-----RCDGVDDCKLSDELGC-----VRPDMDSLKIKYSSGHOMLPICSS 247
 Db 70 FIPRKQCDGELDCPLGEDEBHCYKSPFEGPAVAVRLSKDRSTLQVLDATGNMFSACFD 129
 Qy 248 NMNDYSEKTCQOOLGFESAHRTTEVAHRDPANSESIILRYNSTIOESLHRSRCPQRYISL 307

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLGKMPQVSLRYDGAHLGSGSLSGDWLTAHCF-PPERNRVLSRMVFAGA 59
QY 381 ---SNLHQLPEAASIAETIINSNY-----TDEEDVDIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIQVCLP 117
QY 432 MHGQFSLNETCMTGFKTRETDTKTSPLREVOVNLIDFKKNDVLYVDSYLTPRMC 491
DB 118 AAGQALVDGKICTYTGWNTQYGGQAG-VLQEARVPIISNDVCGADPYGNQIKPKMFC 176
QY 492 AGDLRGGRDSCQSGSGPLVE---QNNRWYLAGVTSWGTGCGGRNRPVYTKVTEVLP 547
DB 177 AGYEGGIDACQDSGGPFVCEDSISRTPRWRLCGIVSWGTCALAQKRGVYTKVSDPRE 236
QY 548 WYSKMES 555
DB 237 WIFQAIKT 244

RESULT 9
US-11-254-185-38
; Sequence 38, Application US/11254185
; Publication No. US20060099625A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; APPLICANT: SCHUMAN, MARC
; TITLE OF INVENTION: MT-SP1 SERINE PROTEASE
; FILE REFERENCE: 28644-701.302
; CURRENT APPLICATION NUMBER: US/11/254.185
; CURRENT FILING DATE: 2005-10-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-254-185-38

Query Match 17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLGKMPQVSLRYDGAHLGSGSLSGDWLTAHCF-PPERNRVLSRMVFAGA 59
QY 381 ---SNLHQLPEAASIAETIINSNY-----TDEEDVDIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIQVCLP 117
QY 432 MHGQFSLNETCMTGFKTRETDTKTSPLREVOVNLIDFKKNDVLYVDSYLTPRMC 491
DB 118 AAGQALVDGKICTYTGWNTQYGGQAG-VLQEARVPIISNDVCGADPYGNQIKPKMFC 176
QY 492 AGDLRGGRDSCQSGSGPLVE---QNNRWYLAGVTSWGTGCGGRNRPVYTKVTEVLP 547
DB 177 AGYEGGIDACQDSGGPFVCEDSISRTPRWRLCGIVSWGTCALAQKRGVYTKVSDPRE 236
QY 548 WYSKMES 555
DB 237 WIFQAIKT 244

RESULT 10
US-11-253-869-6
; Sequence 6, Application US/11253869
; Publication No. US20060104979A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.

APPLICANT: TAKEUCHI, TOSHIHIKO
APPLICANT: SCHUMAN, MARC
TITLE OF INVENTION: MT-SP1 POLYNUCLEOTIDES AND POLYPEPTIDES
FILE REFERENCE: 28644-701.303
CURRENT APPLICATION NUMBER: US/11/253.869
CURRENT FILING DATE: 2005-10-18
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn version 3.3
SEQ ID NO 6
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial
OTHER INFORMATION: Protein fragment/domain
US-11-253-869-6

Query Match 17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLGKMPQVSLRYDGAHLGSGSLSGDWLTAHCF-PPERNRVLSRMVFAGA 59
QY 381 ---SNLHQLPEAASIAETIINSNY-----TDEEDVDIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIQVCLP 117
QY 432 MHGQFSLNETCMTGFKTRETDTKTSPLREVOVNLIDFKKNDVLYVDSYLTPRMC 491
DB 118 AAGQALVDGKICTYTGWNTQYGGQAG-VLQEARVPIISNDVCGADPYGNQIKPKMFC 176
QY 492 AGDLRGGRDSCQSGSGPLVE---QNNRWYLAGVTSWGTGCGGRNRPVYTKVTEVLP 547
DB 177 AGYEGGIDACQDSGGPFVCEDSISRTPRWRLCGIVSWGTCALAQKRGVYTKVSDPRE 236
QY 548 WYSKMES 555
DB 237 WIFQAIKT 244

RESULT 11
US-11-253-869-38
; Sequence 38, Application US/11253869
; Publication No. US20060104979A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; APPLICANT: SCHUMAN, MARC
; TITLE OF INVENTION: MT-SP1 POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 28644-701.303
; CURRENT APPLICATION NUMBER: US/11/253.869
; CURRENT FILING DATE: 2005-10-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-253-869-38

Query Match 17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLGKMPQVSLRYDGAHLGSGSLSGDWLTAHCF-PPERNRVLSRMVFAGA 59
QY 381 ---SNLHQLPEAASIAETIINSNY-----TDEEDVDIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIQVCLP 117

Qy 432 MHGQFSINETCMTGFGKTRTBDKTSPLREVOVNLIDFKKNDYLVYDSYLTTPRMK 491
Db 118 AAGQALVQDKICTVGMWNTGYGQAG-VLQEARVPIISNDVCGADPFYQGIKPKHC 176
Qy 432 AGDLRGSDSCGDSGGPLVCE---QNNRWYLAGVTSWGTCCGGRNPGVYTKYTEVLP 547
Db 177 AGYPRGGIDACGSDGSPVCEBDSISRTPRMRLCGIVSWGTCALAKPGVYTKVSDPRE 236
Qy 548 WYSPKMS 555
Db 237 WIFQAIKT 244

RESULT 12
US-10-196-749-320

/ Sequence 320, Application US/10196749
/ Publication No. US20060094864A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3430R1C340
/ CURRENT FILING DATE: US/10/196,749
/ PRIOR FILING DATE: 2002-07-16
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 320
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-196-749-320

Query Match 17.4%, Score 520.5, DB 6, Length 423;
Best Local Similarity 39.6%, Pred. No. 1e-25;
Matches 109, Conservative 41, Mismatches 112, Indels 13, Gaps 5;

Qy 219 NSFSLRNSTIGSLHRSBPQRYISLQSHCGLRMTGRIYCGALASDKMPWQVSL 338
Db 158 HSKIKIKINKETDYSYLNHCCGTRSKTL-----GQSLRIYGTVEBGEWPMQASL 209
Qy 339 HFGTHICGTLIDQAWYLTAAHCFVYREKVLBGWKYAGTSNHLQPEASIAIEIIN 398
Db 210 QMDGSHRCATLINATWLSAAHCF--TTYKNPARWTASFGVT-IKPSKMRGLRRIIVH 266

Qy 399 SNYTDDEDDYDIALMRSLKPLTSLAHIPACLPMHGQFSINETCMTGFGKTRTBDKT 458
Db 267 EKYKHPSHDYDISLAELSSPVYTNAAVRVCLPDASVEFGPDVNFVYGFGLK-NDGYS 325
Qy 459 SPLREVOVNLIDFKKNDYLVYDSYLTTPRMKCGDLRGSDSCGDSGGPLV-CEQNNR 517
Db 326 QNHLRQAVTLLIDATTCNEPQAYNDATTPRMLCAGSLGKTDACGSDGGLVSSDARDI 385
Qy 518 WYLAGVTSWGTCCGGRNPGVYTKYTEVLPWYISK 552
Db 386 WYLAGVTSWGBDECAKPNKPGVYTRVYALRDWITSK 420

RESULT 13
US-11-101-316-106

/ Sequence 106, Application US/11101316
/ Publication No. US20060099657A1
/ GENERAL INFORMATION:
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
/ TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
/ FILE REFERENCE: P3230R1C17C1
/ CURRENT APPLICATION NUMBER: US/11/101,316
/ PRIOR FILING DATE: 2005-04-06
/ PRIOR APPLICATION NUMBER: 10/063526
/ PRIOR FILING DATE: 2002-05-03
/ PRIOR APPLICATION NUMBER: 10/006867
/ PRIOR FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: PCT/US00/23328
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: 09/380137
/ PRIOR FILING DATE: 1999-08-25
/ PRIOR APPLICATION NUMBER: PCT/US99/12252
/ PRIOR FILING DATE: 1999-06-02
/ PRIOR APPLICATION NUMBER: 60/087759
/ PRIOR FILING DATE: 1998-06-02
/ NUMBER OF SEQ ID NOS: 170
/ SEQ ID NO 106
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-11-101-316-106

Query Match 17.4%, Score 520.5, DB 7, Length 423;
Best Local Similarity 39.6%, Pred. No. 1e-25;
Matches 109, Conservative 41, Mismatches 112, Indels 13, Gaps 5;

Qy 279 NSFSLRNSTIGSLHRSBPQRYISLQSHCGLRMTGRIYCGALASDKMPWQVSL 338
Db 158 HSKIKIKINKETDYSYLNHCCGTRSKTL-----GQSLRIYGTVEBGEWPMQASL 209
Qy 339 HFGTHICGTLIDQAWYLTAAHCFVYREKVLBGWKYAGTSNHLQPEASIAIEIIN 398
Db 210 QMDGSHRCATLINATWLSAAHCF--TTYKNPARWTASFGVT-IKPSKMRGLRRIIVH 266
Qy 399 SNYTDDEDDYDIALMRSLKPLTSLAHIPACLPMHGQFSINETCMTGFGKTRTBDKT 458
Db 267 EKYKHPSHDYDISLAELSSPVYTNAAVRVCLPDASVEFGPDVNFVYGFGLK-NDGYS 325
Qy 459 SPLREVOVNLIDFKKNDYLVYDSYLTTPRMKCGDLRGSDSCGDSGGPLV-CEQNNR 517
Db 326 QNHLRQAVTLLIDATTCNEPQAYNDATTPRMLCAGSLGKTDACGSDGGLVSSDARDI 385
Qy 518 WYLAGVTSWGTCCGGRNPGVYTKYTEVLPWYISK 552
Db 386 WYLAGVTSWGBDECAKPNKPGVYTRVYALRDWITSK 420


```
RESULT 14
US-11-376-673-106
; Sequence 106, Application US/11376673
; Publication No. US20060160186A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
; FILE REFERENCE: P3230R1C165C
; CURRENT APPLICATION NUMBER: US/11/376,673
; PRIOR FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: 10/063742
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 106
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-376-673-106

Query Match      17.4%; Score 520.5; DB 7; Length 423;
Best Local Similarity 39.6%; Pred. No. 1e-25;
Matches 109; Conservative 41; Mismatches 112; Indels 13; Gaps 5;

QY 279 NSFSILRYNSTIQLSLHRSRCPQRYISLQCSHGLRAMTSRIVGALASDSKMPWQVSL 338
Db 158 HSVIKIKINKETPDSYLNHCCTGRSKTL-----QCSLRIVGTEVEGEWPMQASTL 209

QY 339 HFGTHICGGTLIDAQWVLTAAHCFEVTREKYLEGKMYAITSNHLQPEAASIAETIIN 398
Db 210 QMDSHRCGATLINAYMTVLSAHCFT--TTYKNPARMTASFGVT--IKSKMKRGLRRIIVH 266

QY 399 SNYDEEDDYDIALMRSLKPLTLSAHIHPACLPMHGQTFSLNETCWIYFGKTRTDDKT 458
Db 267 EKYGHPSHDYDISLAELSSPVPTNAVHRVCLPDASIEFGQDWMFVTFGALK-NDGYS 325

QY 459 SPFLREVQVNLIDPKKCNLYLVDSYLTFRMMACADLGRDSCQDGGGPLYV-CEQNNR 517
Db 326 QNHRLQAQVTLIDATTCTCEPQAYNDATFRMLCAGSLEGKTDACQDGGGPLYVSSDARDI 385

QY 518 WYLAGVTSWGTGCGQRNKPQYTYKTEVLPWYISK 552
Db 386 WYLAGIVSWGDECAKPNKPQYTYKTEVLPWYISK 420

RESULT 15
US-10-530-187-242
; Sequence 242, Application US/10530187
; Publication No. US20060183120A1
; GENERAL INFORMATION:
; APPLICANT: THE Bin Tean
; APPLICANT: TAKAHASHI, Masayuki
; TITLE OF INVENTION: Molecular Subclassification of Kidney Tumors and the Discovery of
; FILE REFERENCE: 28927-0014
; CURRENT APPLICATION NUMBER: US/10/530,187
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US2003/031476
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: US 60/415,775
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 242
; LENGTH: 853
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-530-187-242

Query Match      17.3%; Score 518.5; DB 6; Length 853;
Best Local Similarity 32.5%; Pred. No. 2.8e-25;
Matches 135; Conservative 50; Mismatches 137; Indels 93; Gaps 17;

QY 199 CPKIAVCDGVNDCKLSDEICVRFPMWKSLSLKIYSGSSHQ-----WLPICS 246
Db 462 CIRKELRCDDGWADCTDHSDELNC-----SCDAGHQFTCKXKFKCKPLFW--VCD 507

QY 247 S--NMNDSYSEKTC-----QQLGFSAHRTTEVAHRFPANFSILR 285
Db 508 SYNDGNSDBEQGSCPAQTFRCNKGKLSKQCNKNDGCDGSDSDASCPKVVVYCTK 567

QY 286 YNSTIQESLHRS-----ECPQRYISLQCS-----HCGLRAMT--GRIVGALASDSK 331
Db 568 HTYRCCLNGCLSKSKNPECDGKE---DCSDGSDKDCDCGLRSFTTRQARVVGTDADGCE 623

QY 332 WPMQVSLH-FQTHICGCTLIDAQWVLTAAHCFEVTREKYLEG-----WKYVAGTSN 382
Db 624 WPMQVSLHAGQGHICGASLISPMWLVSAHCYIDR-----GFRYSDPTQWTAFLG--- 675

QY 383 LHQLEPAAS-----IAETIINSNTDEEDDYDIALMRSLKPLTLSAHIHPACLPMHGQ 435
Db 676 LHDQSRAPQVQDERRLRRIISHPFNDFTFDYDIALLELEKPAEYSSMWRPCLPDASH 735

QY 436 TFSLETWCWITFGKTRTDDKTSPFLREVQVNLIDPKKCNLYLVDSYLTFRMMACADL 495
Db 736 VFPAGKAIWYTGWHT--QYGTGALILQKGEIRIYNOTTCENL--PQGITFRMMACGFL 792

QY 496 RGRDSCQDGGGPLYV-CEQNNRWYLAGVTSWGTGCGQRNKPQYTYKTEVLPWYISK 549
Db 793 SGVYDSCQDGGGPLYVSSDARDIRIFQAGVSWGDECAQQRNKPQYTYKTEVLPWYISK 847
```

Search completed: September 16, 2006, 03:29:58
Job time : 36 secs

This Page Blank (uspio)